

# Improving the Dynamics of Aid

## Towards More Predictable Budget Support

Benn Eifert and Alan Gelb<sup>1</sup>

World Bank Policy Research Working Paper 3732, October 2005

*The Policy Research Working Paper Series disseminates the findings of work in progress to encourage the exchange of ideas about development issues. An objective of the series is to get the findings out quickly, even if the presentations are less than fully polished. The papers carry the names of the authors and should be cited accordingly. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the view of the World Bank, its Executive Directors, or the countries they represent. Policy Research Working Papers are available online at <http://econ.worldbank.org>.*

---

<sup>1</sup> Office of the Vice President, Development Economics, World Bank.

## Abstract

This paper considers approaches towards improving the predictability of aid to low income countries, with a special focus on budget support. In order to accelerate progress towards the Millennium Development Goals, the donor community is increasing aid flows while pushing for more coordination and tighter performance-based selectivity. However, these factors may increase the unpredictability of aid, from current levels, which are already high enough to impose significant costs. Predictability is a particular challenge in the area of budget support, which will continue to increase in importance as aid is sought to underpin longer-term recurrent spending commitments. Budget support reduces transactions costs and drains on capacity, but it tends to be more vulnerable to fluctuations than multi-year project support. Poor predictability raises the threat of a low-level equilibrium: countries, budgeting prudently within a medium-term fiscal framework, will discount commitments; donors will see few funding gaps, so pledges will fall. With some countries discounting aid commitments in formulating budgets, some already see signs of this happening.

To improve predictability, donors must extend their funding horizons. However, even if this can be done, several major issues will remain at country level.

First, how can countries deal with residual short-run volatility of disbursements relative to commitments? The paper demonstrates that simple saving and spending rules built around a buffer reserve fund of 2-4 months of import cover can buffer public spending against short-run fluctuations in aid disbursements.

Second, can donors lengthen commitment horizons to individual developing countries without excessive risk of misallocating aid? We develop a simple model of aid allocation and efficiency, and estimate the deadweight losses that might follow from pre-committing aid over several years, rather than one year at a time, on the basis of an initial performance assessment. We find that efficiency losses can be modest under a system of “flexible pre-commitment”, and that very large improvements in predictability from current levels can be reconciled with performance-based aid allocation.

Third, within a country’s overall aid envelope, how should donors set the shares of project aid and budget support? Arguing that budget support is in part an investment in country systems, we suggest that changes in performance, in addition to levels, can be useful indicators of the payoff to budget support. Triggers and prior actions can be used as interim indicators of improving performance. But aid flows cannot respond too sharply to such indicators while keeping a reasonably stable resource flow, and interim assessments need to be complemented by deeper performance assessments, including through survey methods, every several years. We suggest guidelines for the responsiveness of budget support to performance that also take into account the likely degree of uncertainty in measuring performance.

Finally, the paper considers the other main approach to budget support, the output or outcome-driven approach of the European Union. We do not argue for one approach or the other; indeed, we believe policy-based and outcome-based approaches to be complementary. We focus on a current difficulty with output-based support, how to set targets for social indicators and service delivery standards. Using previous studies and quantile regressions, we illustrate how historical experience can help frame forward-looking goals and improve the foundation for results-based assistance, for gross primary enrollment and infant mortality indicators.

The paper concludes that many of these issues can be addressed. Simple spending and savings rules built around a buffer reserve fund of 2–4 months of imports can help smooth public spending. Aid can be pre-committed several years ahead with only small efficiency losses, using a strategy of “flexible pre-commitment.” Guidelines can be set to limit the volatility of budget support while keeping it performance-based, and past experience can be used more systematically to develop “outcome” norms to better guide aid allocation.

## 1 Introduction

Low-income countries (LICs) face many sources of instability. Their concentrated primary economies render them particularly vulnerable to climate and terms-of-trade shocks, which they are less well equipped to buffer compared with higher-income countries. LICs often suffer from political crises, including destabilizing regime changes. While LICs are less integrated into capital markets than middle-income countries (MICs), volatile financial flows in the form of unpredictable aid disbursements constitute another source of macroeconomic and fiscal instability for LICs. Like private capital flows, changes in aid can occur exogenously or in response to perceived changes in governance and economic management. Aid cannot be completely stable if it is to respond to performance and also provide disaster relief, but studies suggest that predictability is surprisingly low, even for non-emergency flows. They also suggest that aid has tended to be mildly pro-cyclical. Unpredictability and pro-cyclicality impose high costs on vulnerable, aid-dependent economies.

This paper considers approaches towards improving the predictability of aid, in particular program aid delivered in the form of budget support. With the donor community moving towards increased aid flows, greater donor coordination, and increasing selectivity, predictability is becoming more urgent, since these factors are likely to increase the volatility of aid flows. The predictability of program aid is particularly important, as donors have been shifting towards budget or sector support to reduce the transactions costs and drains on capacity caused by the need to implement large numbers of projects. This promises to reduce the transactions costs of aid, and offers an incentive to strengthen country systems. However, program aid tends to be more vulnerable to fluctuations than project support, which is usually committed up front and disbursed on a multi-year basis. The problem of predictability will only become more serious as aid is sought to underpin longer-term recurrent spending commitments, such as recruiting teachers and increasing the pay of medical personnel.

If aid predictability is not improved, countries moving from uncoordinated project support to coordinated program support may therefore see any decline in transactions costs offset by an increase in the costs of volatility. This increases the danger of slipping into a low-level equilibrium: countries, budgeting prudently within a sound medium-term fiscal framework, will discount pledges of assistance; donors will see few funding gaps, this in turn will cause pledges and commitments to fall. With many countries reportedly discounting aid commitments heavily in formulating budgets, some already see signs of this happening.

Several problems must be solved to improve aid predictability. Progress is, of course, needed to extend the funding horizon on the donor side. This includes finding ways to make multi-year commitments, possibly underpinned by mechanisms like the proposed International Financing Facility, or moving to the three-year framework of IDA or the Millennium Challenge Account, and extending these for longer periods. More streamlined processes could help disbursements to better reflect pledges and commitments, while strengthening the annual review and aid programming cycle at the country level can also improve predictability<sup>2</sup>.

---

<sup>2</sup> For recent approaches to improving the quality of aid, see OECD/DAC 2004,2005.

However, even if progress is made on these fronts, four major issues will remain:

- First, how can countries deal with residual “exogenous” short-run volatility of disbursements relative to commitments?
- Second, can donors lengthen commitment horizons from one year to several, without excessive risk of misallocating aid between countries?
- Third, within the overall aid envelope, how can donors strike an appropriate balance between responding to levels of performance versus changes in performance when setting the shares of project aid and program or budget support?
- Fourth, how can historical experience be drawn on to derive output or outcome targets which can be used to determine the level of assistance?

This paper addresses these topics, with emphasis on the dynamics of aid flows.

Section 2 summarizes the basic stylized facts emerging from the literatures on volatility and aid instability, focusing on the magnitude and the determinants of past aid volatility, the degree of pro-cyclicality in disbursements, and the likely economic costs of volatility and uncertainty. Sections 3 – 6 turn to the four questions in a forward-looking framework.

Section 3 addresses the exogenous component of aid volatility. We consider how to integrate aid volatility into fiscal programming and reserve management in order to buffer development spending against exogenous aid shocks and allow time for donors to compensate for under- or over-disbursement relative to commitments. This requires appropriate rules for spending and reserve management. Simulations of a simple system suggest that for most countries a buffer on the order of 3 months of import cover can enable a smoothly functioning corrective feedback loop. Naturally this is not the end of the story: domestic revenue shocks might need to be buffered also, and pooling buffer reserves across countries could economize on the overall stock of reserves needed. But estimates of need and rules for access would still need to be considered at the country level. Another issue, touched on later in the paper, is how aid could be made more counter-cyclical by shortening information lags on country-level impacts of global trends.

Section 4 considers the tradeoff between predictability and performance-driven aid. Can the donors’ commitment horizons be lengthened to several years, or is this too risky given drift in the quality of policies and governance over time? Taking IDA as an example, and using only very general assumptions, we estimate the deadweight losses that would have resulted over 1999-2003 from pre-committing allocations to countries for periods of 5 years rather than annually. Losses are modest in most cases, and become very small with a “flexible commitment” rule where changes only kick in when governments’ performance ratings rise or fall substantially. Indeed, the losses are smaller than those due to the likely error in measuring performance. This suggests that aid can safely be committed with a multi-year horizon, subject to provisions related to “catastrophic” performance failure or “quantum” improvements.

Section 5 turns to the place of budget support within the overall aid envelope and the role of conditionality in setting the breakdown of total assistance into budget support and project support. Taking IDA as an example again, why should prior actions and triggers be used to determine this breakdown? Are they just a hangover from the previous conditionality of structural adjustment programs? Why not just set budget support as some proportion of the

overall allocation, perhaps determined by a performance benchmark (such as the World Bank's Country Policy and Institutional Assessment [CPIA])? We suggest an explanation for the continued use of triggers: that budget support is also an investment in improving country systems.. Triggers and prior actions signal the return to this investment by showing countries' commitment to improve their performance in return for donors agreeing to channel funds through their still-fragile country systems.

The question then is how sharply budget support should respond to **changes in performance** versus **levels of performance**. We argue for establishing a code of good practice where core budget support is stable for several years, complemented by graduated performance-levels-based changes in the share of aid provided as budget support and deep, periodic, assessments of changes in budget management and the efficiency of service delivery. In interim years, prior actions or other indicators of performance change could trigger modest incentive payments, along the lines of front-loading or back-loading in a typical Country Assistance Strategy (CAS). "Catastrophic" breakdowns in financial management, budget discipline or macro-management which compromise the effectiveness of budget support would warrant sharp responses, both for budget support and for overall assistance. The graduated response of course raises the need for donors to be able to decide where macroeconomic slippage is "catastrophic" rather than rely only on an on-off judgment from the IMF.

Finally, in Section 6 the paper considers the other main approach to budget support, the outcomes-driven approach of the EU. We do not debate whether budget support is better determined by assessments of policies and institutions (as in the current Bank approach) or service-delivery or other outcome targets. Indeed, these approaches should be complementary; tracking outcomes is essential for assessing policies, and understanding the causes of the outcomes, including the underlying policies, is essential for using targets intelligently to guide resource use. What is relevant to this paper is that the EU framework provides a useful approach to the tradeoff between predictability and performance. However, while the final relationship between outputs and disbursements in EU programs is clear, the determination of the size of the fixed tranche and the crucial question of how to set the targets for the indicators remains less well determined. Countries setting "stretch goals" will be more likely to be penalized than countries with more modest objectives, so that setting targets requires some comparative reference<sup>3</sup>. We consider how service delivery and progress norms could be derived from the historical experience of developing countries to help interpret a country's goals and improve the foundation for outcome-based assistance. This is illustrated using the examples of gross primary enrollment and infant mortality. The results suggest that goal-setting in EU programs is indeed ambitious.

One important issue for the paper is the implications of performance measurement errors for fine-tuning support. CPIA ratings and outcome data are subject to error, and even if not very large, these can obscure the true size of changes over the short-run. Optimal control theory suggests to dampen responses when noise/signal ratios are high; this provides an additional argument for not leveraging modest changes in performance into major swings in assistance.

Section 7 concludes, with a summary of implications.

---

<sup>3</sup> One of the strengths of the CPIA system is its explicit attempt to be comparative.

## 2 Aid Instability: Some Stylized Facts

The issue of aid instability encompasses a number of dimensions. Volatility and unpredictability often go together, but volatility is not always unpredictable (e.g. reliable provision of emergency food aid). Volatility can be endogenous (performance-related) or exogenous. Unpredictability can be very short-term (intra-year), short-medium-term (over 1-3 years) or long-term in nature. Aid instability can have microeconomic and microeconomic implications, and can also influence institutions by changing the rules and incentives affecting behavior. Not all studies of aid instability cover all dimensions, but together they offer a reasonably clear picture of experience.

**How volatile is aid?** In the aggregate, aid looks quite stable: Osei, Morrissey and Lensink (2002) suggest that the total sum of ODA to low-income countries was less variable than the total sum of either FDI or other private capital flows over 1970-97. However, the volatility of aid flows to individual countries is much higher, and it also varies significantly, with aid to better-managed countries often less volatile. Standard deviations around time trends and autoregressive forecasts are generally smaller than the coefficients of variation -- for instance, in Indonesia the coefficient of variation is 78% of the mean aid flow, but the standard deviation around an autoregressive forecast is only 19%. Predictability based on past trends is therefore not quite as poor as measurements of overall volatility suggest, but both are substantial: the median standard deviation of disbursements is 37 percent of the mean, and it is 34 percent relative to an autoregressive predictor of the expected level. Bulir and Hamann (2003, 2005) find that the volatility of aid is very large, with coefficients of variation usually in the range of 40-60% of the mean aid flow, and exceeds that of fiscal revenues. They find that volatility rises with the level of aid dependence -- the median volatility of aid is almost five times as high in the 33 countries with aid:revenue ratios of 50% or more than in their entire sample of 72 countries. They also find that aid volatility tends to be higher in countries which also have higher revenue volatility, and that program aid tends to be more volatile than project aid. The seriousness of the problem is underlined by another finding: in most aid-dependent countries, commitments convey no more information on future disbursements than do past disbursements. This is astonishing, given the importance placed on commitments as an input into medium-term fiscal frameworks.<sup>4</sup>

**Is aid volatility declining?** Recent reassessments suggest that, despite efforts to harmonize aid around the PRSP process and improve predictability, there has not been progress on the latter front. Osei et al, Gabriele et al (2000) and by the more recent review of Bulir and Hamann (2005) all find evidence that volatility of all types of flows to both low- and middle-income countries has increased in the 1990s. As commitments have risen, disbursements have fallen behind, to about 60% of commitments. Despite the fact that aid patterns in the post-Cold-War era should in principle have become more development-friendly, the title of the Bulir-Hamann (2005) paper is suggestive: "Volatility of ODA: From the Frying Pan into the Fire."

---

<sup>4</sup> In a simple autoregressive model of disbursements, coefficient on commitments is significant at 5% level in only one-third of all countries. In these cases the coefficient is around 0.4 on average and 0.3 at the median, suggesting that even here commitments contain only partial information. The only countries for which the coefficient is larger than 0.7 are Argentina, Mexico, Panama, Turkey and Venezuela, all low aid-dependency countries. In the high-aid-dependency countries, the coefficients on commitments were significant in only one-fifth of cases

**How much of aid volatility is “exogenous” and how much is directly performance-related?**

It is difficult to disentangle these factors because donors are not always clear about their reasons for changing levels of support. Performance-related factors influence project disbursements but their impact on overall volatility is unclear. Bulir and Hamann (2003) show that, as might be expected, IMF program status does not affect project aid. However, it does influence program aid -- off-track countries receive 33% of commitments, compared to 75% for on-track countries. This result crudely suggests that about 70% of the volatility of program aid might be due to performance-related issues, and 30% to other factors. A 2005 assessment of donor's views by the SPA Budget Support Working Group indicates that 40% of non-disbursements were considered to be due to failure to meet policy conditionality, 25% to recipient governments' delays in meeting administrative conditions, 29% to administrative problems on the donor side, 4% to political problems on the donors side and 2% to other factors. A rough rule of thumb might therefore be that around half of the volatility in program assistance might be performance-related while half reflects administrative delays and other exogenous factors.

**Is aid pro-cyclical or counter-cyclical?** Several studies, including Bulir and Hamann (2003), Gemmell and McGillivray (1998), and Pallage and Robe (2001), find that aid is mildly pro-cyclical, tending to move in the same direction as GDP and domestic revenue. Various factors have been put forward to account for this. Some have hypothesized that pro-cyclicality arises from correlated business cycles in the North and the South which cause aid budgets to tighten during downturns, but this proposition finds no empirical support (Pallage and Robe (2001)). A more interesting argument revolves around the complementary role of counterpart funds. Pallage and Robe (2003) develop a model in which the pro-cyclicality of aid results from the divergence of preferences between donors and recipients. If recipient governments value both the social returns of projects and some other output (for instance, disproportionate benefits to favored constituencies), it is rational for donors to require counterpart funds in order to reduce the fungibility of resources to “low-return” projects that would otherwise accompany increased aid. Shocks or economic downturn that cut counterpart funds then cause aid flows to projects to fall, an especially likely outcome in the poorest and most vulnerable countries. Another donor-discretion-based hypothesis which has not yet been explored concerns the imperfect observability of policies or “effort” on the part of recipient governments. If economic performance is a function of both governments' overall development efforts and other factors, and if “effort” is difficult to observe, donors will associate an observed improvement (deterioration) in economic performance with an improvement (deterioration) in effort, leading to pro-cyclical aid flows. Countries doing well “must be doing something right” and therefore get more aid. While facilities have been developed to provide counter-cyclical finance (e.g. CFF and STABEX), studies suggest that their effectiveness has been undermined by decision delays, slow procedures, and lags in information (particularly about the coming impact of newly-identified terms-of-trade shocks).<sup>5</sup>

---

<sup>5</sup> Rapid counter-cyclical disbursement mechanisms are difficult to implement, both because of performance criteria and because shocks are not always easy to recognize as such when they occur. The EU Stabex and Sysmin programs operated between 1975-2000 and disbursed Euro 6.1 billion in total. Stabex increased conditionality and payment-justification schemes in the wake of the collapse of most domestic micro-stabilization schemes in an attempt to ensure that benefits reached farmers, but this slowed disbursements badly, such that assistance actually became pro-cyclical, amplifying the effects of commodity cycles (Hermann, Burger and Smit 1990). The EU Flex scheme began in 2000, aiming at fast-disbursing aid in response to two conditions: a) losses of overall export earnings from goods by more than 10%; b) 10% worsening in the programmed public deficit. So far six countries have been declared

**Do political explanations of aid allocations account for aid volatility?** Cross-country evidence suggests weak performance-based patterns in aid allocation. Neumayer (2003a) investigates a broad set of governance variables, finding that none show a consistent pattern of significance across the group of donors, though democracy, human rights, corruption, military expenditure, rule of law and regulatory burden are all significant for some donors. Alesina and Dollar (2000) note that, for major bilateral donors, political-interest variables like colonial history and voting similarity in the UN General Assembly are more important determinants of aid allocation than governance and policy. Andersen, Hansen and Markussen (2004) find that voting similarity to the US in the UN is a significant determinant of World Bank *commitments* (not disbursements) during 1991-2000, while Andersen, Harr and Tarp (2004) also find that ‘political concessions’<sup>6</sup> in UN votes which the US State Department identifies as important are significant determinants of IMF loans. Interestingly, Neumayer (2003b) finds that several UN agencies, as well as the African and Inter-American Development Banks, tend to counteract certain biases of bilateral donors, giving less to former colonies of large donor countries. Dollar and Levin (2004) find that, for many donors, aid has become increasingly selective on the basis of policies, governance and poverty over the 1990s: their top performers are IDA, IMF, the UK and the Scandinavian countries; the exceptions unfortunately include some of the largest donors, such as the US and France. There is somewhat less work on changes in aid over time, but Alesina and Dollar (2002) find that, while donors do respond to democratization with large surges in aid, the same pattern does not hold for economic reform.<sup>7</sup>

None of these studies focuses on the predictability of aid, and the mechanisms which they document are unlikely to cause short-run unpredictability. Most political interest variables are slow-moving, and aid fluctuations based on votes in the UN General Assembly are a predictable element of international relations. Similarly, since movements into and out of democracy are uncommon, it is unlikely that they constitute a major cause of aid instability in most countries. Nevertheless, the studies point to the continued medium-run vulnerability of aid flows to political factors and the significance of explicit or implicit political conditionality. This becomes more important in the context of efforts to provide larger, and more stable, aid flows to support efforts to meet the MDGs, since the share of aid in such scenarios rises to over half (and sometimes up to two-thirds) of overall budgets.

**Is volatility really costly?** Most studies on the impact of volatility have been carried out at macroeconomic level, using growth or indicators of consumption and risk aversion to estimate the impact. Whereas the costs seem to be modest for diversified, high-income countries, those

---

eligible for support and Euro 36 million has been disbursed. The IMF has used the compensatory financing facility (CFF) 344 times in the 1970s and early 1980s, mostly for export revenue shortfalls. However, it has not been used since its 2000 review. It was not concessional, so not particularly attractive for LICs. IMF has recently proposed changes to create more consistent responses to shocks in LICs, including augmenting existing on-track PRGFs and subsidizing some GRA facilities in countries that do not have PRGFs or are off-track. The Fund is recommending that for the weakest countries, its role should be to draw attention to needs, cooperating with donors capable of providing grants. The Bank has occasionally provided supplemental financing to countries already on a multi-year adjustment program and experiencing adverse trade shocks.

<sup>6</sup> These are measured as movement from a country’s “bliss point” towards the US position

<sup>7</sup> Consistent with literature that finds that aid has not “bought” reform, the authors also find no tendency for policy reform to follow surges in finance.



for poor undiversified economies with greater rigidities and tight liquidity constraints appear to be considerable: the welfare costs of output instability in Sub-Saharan-Africa might be 15-20 times those in the US (Pallage and Robe 2003). Pallage and Robe (2003) suggest that moderately risk-averse consumers in Africa might be willing to trade off an extra one percent annual improvement in living standards to eliminate volatility. Collier and Dehn (2001) show that a 40% negative price shock to exports reduces growth by 1.5% per year, for a 5.5% loss over a 4 year period. These results echo those of simulations undertaken for resource-rich economies seeking to manage volatile natural resource rents: Gelb (1988) shows that the costs of volatility induced by errors in projecting oil prices can easily overwhelm the benefits of export-driven spending booms, even assuming optimal savings and spending profiles along the projected oil price trajectory. Loss functions are not symmetric: the losses from big adverse shocks exceed the gains from big positive shocks.

There are no comparable estimates of the efficiency cost of unstable budgetary revenues because of lack of consistent data, but anecdotal evidence suggests that they are very large. On the upside, rapid spending increases are often wasteful; on the downside, governments lack the recurrent resources needed to complement capital investments and to complete unfinished projects. Spending rigidities, especially salaries, crowd out essential non-salary inputs; cash-strapped governments turn to thin domestic financial markets, crowding out the private sector and sparking high-interest domestic debt spirals that threaten macroeconomic stability. In addition, high volatility in fiscal resources undermines results agreement and accountability mechanisms between donors and governments and among ministries within governments. Kostopoulos (1999) found that that budget instability in Africa was very high: only 45 percent of countries experienced less than 10 percent deviation of aggregate spending from projections; 33 percent experienced deviations of 10-30 percent, while 22 percent experienced deviations from projections of above 30 percent. Because certain expenditures are less compressible than others, deviations are magnified at sector and program level: about half of all spending programs at sector level deviated from projections by 30 percent or more. Unstable aid was of course not the only, or perhaps even the main, cause of deviations. But the results suggest how far the combination of improved budget management and more predictable resource flows needs to improve to provide a stable base for development spending.

### **3 Cushioning Expenditures through Reserve Management and Fiscal Rules**

Reserves represent countries' first line of defense against shocks. The extensive literature on central bank reserve management focuses primarily on middle-income countries and their responses to commodity price shocks and private capital flows. Some studies provide strong arguments for ample reserves: for instance, Aizenman, Lee and Rhee (2004) who study a model of exposure to sovereign risk and downside output risk associated with a costly debt crisis implies a relatively high optimal levels of reserves.<sup>8</sup> General reserve management principles

---

<sup>8</sup> One important caveat is noted by Aizenman and Marion (2004). A greater chance of corruption and opportunistic behavior by future policymakers reduces the demand for international reserves and increases external borrowing, such that a high debt-to-reserves ratio is a symptom of poor governance. They suggest that, in such cases, the traditional policy recommendation to increase international reserve holdings may be welfare-reducing.

can be adapted to particular contexts of instability using specific institutional mechanisms. The most commonly cited in the developing world is Chile's copper revenue stabilization fund<sup>9</sup>.

This section uses a simulation model to study the potential of a parallel mechanism for managing the *exogenous* volatility of aid inflows. We highlight three main points. First, a relatively simple scheme based on a reserve tranche of 50-100% of annual aid-financed spending (2-4 months of import cover in a typical aid-dependent low-income country) can be effective in smoothing expenditure in most periods under a range of levels of aid instability. Second, while our simulated stabilization fund does in some cases go "bankrupt", this usually requires 3-5 years of large negative shocks to aid flows. Countries and donors should have plenty of lead time before a full run-down, enough to organize an emergency response. Third, for instability in the high end of the range we consider, the reserve levels required are significantly higher. With moderate reductions in instability, the necessary reserve cover could be cut substantially.

Our stabilization fund is a simple one. Unplanned deviations from aid-financed spending targets should be kept within a small percentage of target levels. When the reserve buffer exceeds its long-run target, the fund operates in "high mode", protecting against downside shocks more vigorously. When the fund's reserve stock is below its long-run target, the fund operates in "low mode", a more cautious framework which attempts to maintain spending levels while replenishing its reserve stock. One could imagine much more sophisticated mechanisms for the management of aid volatility, but this simple instrument suits our illustrative purposes here.

Suppose that donors agree to finance a recipient government's medium-to-long-term development plan, but their commitments are subject to a stochastic shock. That is,  $a_t = \tilde{a}_t + \theta_t$ , where  $\theta_t$  is potentially serially correlated, e.g.  $\theta_t = \lambda\theta_{t-1} + \varepsilon_t$ , where  $\varepsilon_t \sim N(0, \sigma^2)$ .

Define a target path for aid-financed spending,  $G_t^*$ , which donors commit to finance with aid  $a_t$ , such that  $G_t^* = \tilde{a}_t$ . Denote actual program spending as  $G_t$ . The buffer stock  $S_t$  is designed to smooth  $G_t$  in the face of fluctuations in  $a_t$ .

A simple rule for a buffer mechanism is as follows:

- Target an equilibrium buffer stock level  $S^*$  equal to a share  $s$  of targeted aid-financed spending  $G_t^*$ , such that  $S^* = sG_t^*$ ;
- Designate a program spending floor  $\underline{G}_t$  and spending ceiling  $\overline{G}_t$ ;
- In year  $t$ ,
  - If last period's year-end buffer stock level  $S_{t-1} < S^*$ , then:
    - If  $a_t < \underline{G}_t$ , then  $G_t = \underline{G}_t$ , with the deficit financed by the buffer stock;
    - If  $\underline{G}_t < a_t < 0$ , then  $G_t = a_t^*$ , so no change in the buffer stock;
    - If  $a_t \geq G_t$ , then  $G_t = G_t^*$ , and the surplus replenishes the buffer stock.
  - If  $S_{t-1} > S^*$ , then:
    - If  $a_t > \overline{G}_t$ , then  $G_t = \overline{G}_t$ , and the surplus augments the buffer stock;

---

<sup>9</sup> Together with Chile's Structural Balance Rule, the stabilization fund has enabled the conduct of credible counter-cyclical fiscal policy: see Perry 2002

- If  $0 > a_t > \bar{G}_t$ , then  $G_t = a_t^*$ , so no change in the buffer stock;
- If  $a_t \leq \bar{G}_t$ , then  $G_t = G_t^*$ , with the deficit financed by the buffer stock.
- If  $S_{t-1} = S^*$ , then:
  - $a_t < \underline{G}_t$ , then  $G_t = \underline{G}_t$ , with the deficit financed by the buffer stock;
  - $\underline{G}_t > a_t > \bar{G}_t$ , then  $G_t = a_t^*$ , so no change in the buffer stock;
  - $a_t > \bar{G}_t$ , then  $G_t = \bar{G}_t$ , and the surplus augments the buffer stock.

### *Simulating the Performance of the Buffer Stock*

We randomly generate 1,000 successive observations of  $\theta$ , based on parameter choices discussed below, and then impose the spending rule under its own range of parameters to see what the performance of the buffer fund is. A key metric of performance is the percentage of years in which the buffer fund's balance runs negative. We also present graphs of the buffer fund's balances over time and investigate the characteristics of negative shocks – what does it take to drive the fund into the red, and how predictable would the bankruptcy of the fund be?

The standard deviation of program aid flows from their targets ( $\sigma$ ) is of crucial importance to the viability of buffer mechanisms. Following Bulir and Hamann (2003), we take the endogenous-exogenous breakdown as 70:30, which suggests that an appropriate range of  $\sigma$  to consider for historical exogenous aid shocks is 0.12 – 0.18 of mean aid flows. To this range we add a lower bound (0.10) and an upper bound (0.20) for sensitivity analysis<sup>10</sup>

For the size of the equilibrium buffer stock, we consider the range from 50-100% of the annual aid flow. For an aid-dependent low-income receiving 10% of GDP annually in budget-support aid and with annual imports of close to 40% of GDP (the average for PRSC countries), this corresponds to 2-3 months of import cover, less than their current average cover of almost 5 months imports. The spending floor and spending ceiling determine both the usefulness of the model in smoothing expenditure and the vulnerability of the buffer stock to persistent shocks. The closer the floor and ceiling are to the target expenditure level, the greater the expenditure-smoothing effect, but also the more easily the fund can be depleted by persistent negative shocks to aid flows. For illustrative purposes we use plus and minus five percent of the spending target to designate the floor and ceiling.<sup>11</sup>

Finally, and perhaps most importantly, is the nature of the variability in aid flows. If aid evolves according to a stationary process, such that  $\lambda = 0$  and  $\theta_t$  is purely stochastic, simple buffer stock mechanisms may be quite effective. In the best of all scenarios, if donors offset unreasonably weak aid flows in time  $t$  by increasing aid in  $t+1$  (e.g.  $\lambda < 0$ ), a simple buffer mechanism would be

<sup>10</sup> Note that this may be interpreted as an upper-bound estimate for the share of exogenous causes in the variation of aid, because performance can still vary over a significant range without causing an IMF program to go off-track.

<sup>11</sup> The 5 percent limit is arbitrary, but would represent a large improvement on current conditions.

highly effective in smoothing spending, but the more persistent are aid shocks, the more vulnerable the buffer mechanisms will be.. We consider four values of  $\lambda$ : -0.25, 0, 0.25 and 0.5.<sup>12</sup>

The benefits of the buffer fund are clear from the simulation. While the fund retains positive balances, its operation reduces the standard deviation of actual spending to 2.8 – 3.3 percent of mean spending under all parameter configurations. By definition, spending never falls below its target level by more than five percent.

The problem that must be contained is the vulnerability of the buffer stock to persistent negative shocks. Table 1 shows the percentage of years in which the buffer stock has a negative balance, under a range of parameter values for  $\sigma$  (overall aid volatility levels),  $\lambda$  (autocorrelation in aid shocks), and  $S$  (target stock as share of aid-financed spending). Taking a negative-balance rate of 10 percent of years or below as an adequacy benchmark, if aid shocks are negatively autocorrelated ( $\lambda = -0.25$ ), our smallest fund ( $S = 0.5$ ) can easily dampen anything less than the highest volatility we consider ( $\sigma = 0.20$ ). Serial independence of shocks implies very low vulnerability as well, particularly for the larger funds. The more positively autocorrelated are aid shocks, the more vulnerable is the buffer fund, because bad shocks tend to be persistent. With  $\lambda = 0.5$ , even the largest buffer fund we consider ( $S = 1$ ) can only keep the years of negative balance less than 10% of all years in the mildest volatility environment we consider ( $\sigma = 0.10$ ). For a milder level of autocorrelation ( $\lambda = 0.25$ ), the larger buffer stocks can successfully dampen moderate levels of volatility ( $\sigma = 0.12$  for  $S = 0.75$ ,  $\sigma = 0.15$  for  $S = 1$ ).

To illustrate the functioning of the buffer stocks, Figures 1 – 4 trace the 1,000-year simulations for the smaller fund ( $S = 0.5$ ) in a country receiving 10% of GDP in aid. The illustration shows the middle ranges of volatility ( $\sigma = 0.12, 0.15$ ), across the range of possibilities for autocorrelation. Zero or negative autocorrelation ensures that the fund's balance fluctuates fairly closely on the range of 0 – 10 percent of GDP, only going negative in the case of persistent large negative shocks (eg around year 100), but in the higher-autocorrelation cases the small fund is more vulnerable.

Figure 5 zooms in on the episode of shocks concentrated in the simulation years 21-59, to investigate the “anatomy” of the worst series of shocks generated by the simulation under the strongest autocorrelation ( $\lambda=0.5$ ). In the worst episodes, 25-30 and 54-59, the buffer fund has a reasonably positive balance to begin with, and even the  $\sigma = 0.20$  cases require 3-4 years of sustained large negative shocks before the buffer fund is depleted. This is comforting. It implies that countries and donors will have plenty of lead time before a full run-down of reserves occurs. If consultative group arrangements can work reasonably well, and if there is a clear performance framework to help define when aid shocks are exogenous, countries should be able to avoid severe crises. The better such processes work the lower can be the reserve fund.

Aid-dependent low-income countries, together with donors and the IFIs, may therefore wish to consider more active fiscal programming and reserve-management arrangements in order to reduce the negative effect of exogenous volatility in aid flows. Such instruments can work well as long as volatility does not grow too extreme and negative shocks are not too persistent, and

---

<sup>12</sup> To our knowledge there has been no work which estimates the degree of autocorrelation of aid disbursements relative to commitments. Future research could help to underpin forward-looking plans for volatility management.

they can also underwrite an effective consultative process to correct disbursement levels to better reflect trends in commitments. Our simulation suggests that the upper range of current levels of purely exogenous volatility (equivalent to 18-20% of mean aid flows) would be difficult to buffer effectively. Yet even here it still takes several years of large negative shocks to exhaust a buffer fund, which in principle allows for a compensatory feedback process to work well. Relatively modest decreases in exogenous volatility (to 10-15% of mean aid flows) have high returns in lowering the vulnerability of the buffer fund.

**Table 1. Percentage of years with negative buffer stock balance**

S→	0.5				0.75				1.0			
$\sigma \downarrow$	$\lambda = -0.25$	0	0.25	0.5	$\lambda = -0.25$	0	0.25	0.5	$\lambda = -0.25$	0	0.25	0.5
0.10	1	3	6	24	0	1	3	11	0	0	2	6
0.12	3	6	14	34	0	3	5	24	0	2	3	16
0.15	7	14	26	41	3	5	15	35	2	3	8	28
0.18	9	21	31	42	5	10	25	38	3	5	17	33
0.20	16	30	38	47	7	18	31	43	4	10	24	39

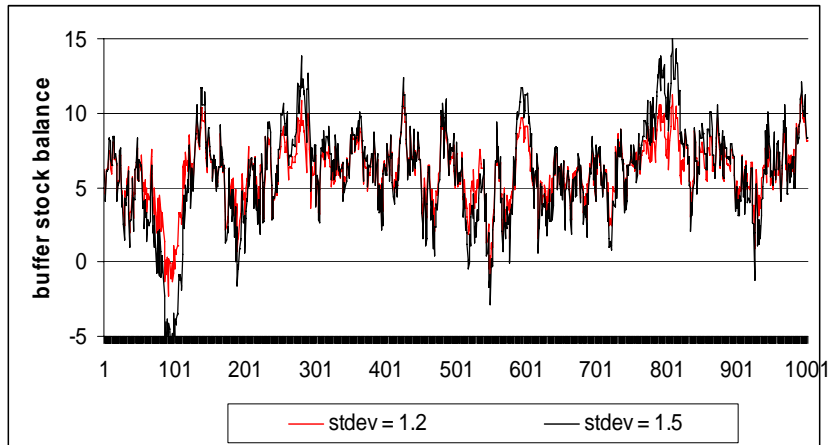
#### 4 How Risky Are Multi-Year Aid Commitments?

The issue of multi-year aid commitments is a controversial one. On one hand, they can introduce a greater degree of stability into fiscal programming. On the other, long-horizon commitments run the risk of over- or under-providing aid in the event of significant changes in country performance. This section lays out an approach to thinking about the efficiency losses from sub-optimal aid allocations. It then simulates this model to ask the following questions: given countries' actual scores on the CPIA (our proxy for performance) over 1999-2003, how large would the efficiency losses from aid misallocation have been if five-year donor programs had been implemented in 1999 under different pre-commitment rules? And, what is the magnitude of the reduction in aid volatility resulting from these rules?

##### *A Simple Model of Aid Effectiveness*

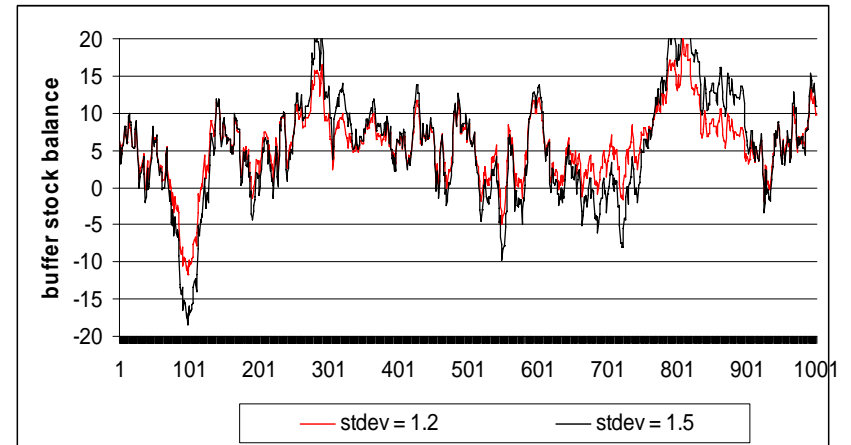
We depart from two main propositions. First, there is some optimal allocation of a given total of aid across countries that depend on poverty levels, population, and a measure of "absorptive capacity" or "performance". Second, the optimal shares of individual countries are independent of total aid. Though not too restrictive, these assumptions do have some implications. One is that aid does not have increasing marginal returns over any range; another is that all countries, even if poorly governed, will still have some aid opportunities with high social returns. The problem in poorly governed countries is that these opportunities run out quickly, and the possibilities for more complex interventions, which depend on domestic management and complementary public sector actions for their effectiveness, are scarce. Such countries therefore

**Figure 1. Buffer Stock Trajectory,  $\lambda = -0.25$ ,  $S = 0.5$**

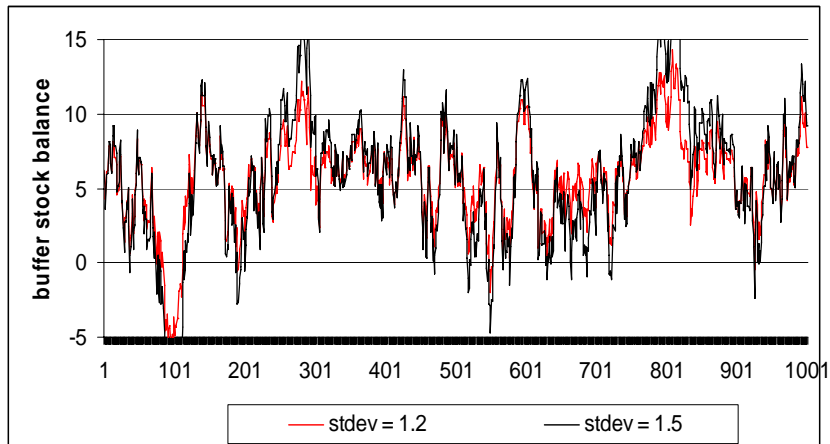


\*Monte-Carlo simulation for 1000 years.

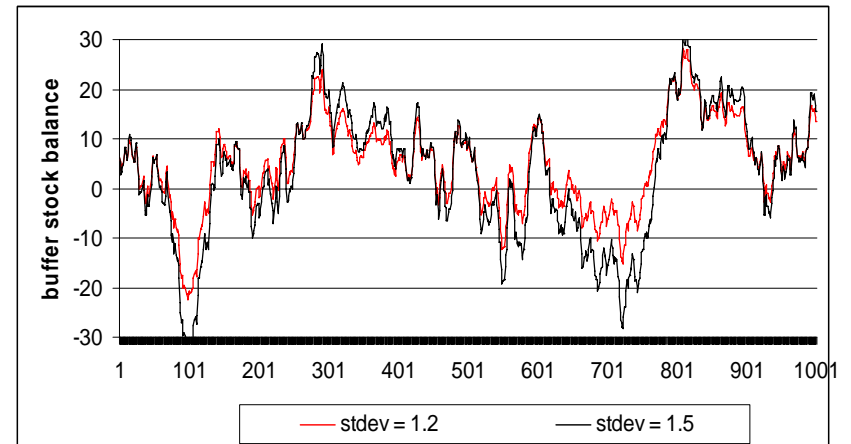
**Figure 3. Buffer Stock Trajectory,  $\lambda = 0.25$ ,  $S = 0.5$**

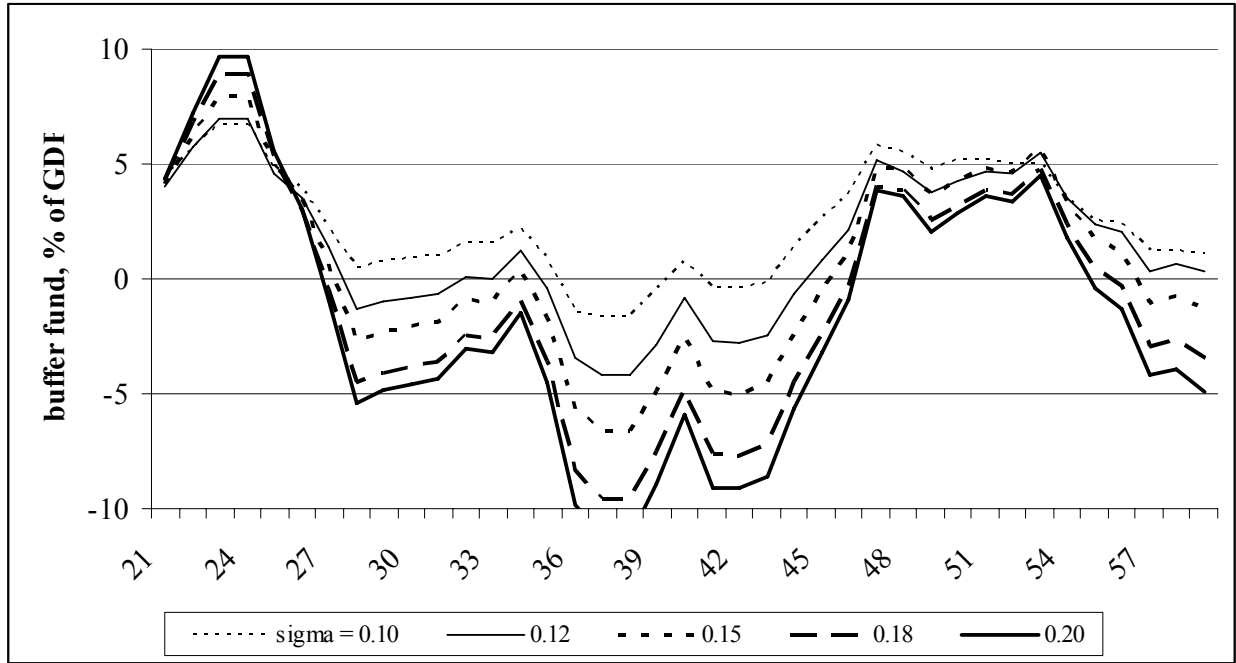


**Figure 2. Buffer Stock Trajectory,  $\lambda = 0$ ,  $S = 0.5$**



**Figure 4. Buffer Stock Trajectory,  $\lambda = 0.5$ ,  $S = 0.5$**



**Figure 5. Anatomy of Major Shocks**

\* $\lambda=0.5$

receive less, but not zero, aid. It follows from the first and second assumptions that the marginal value of aid decreases linearly with the level of aid, whether expressed as aid/head or aid/GDP. A number of studies have estimated the returns to aid using functions that are compatible with the above propositions, including Burnside and Dollar (2000), Collier and Dollar (2002), Denkabe (2003) and Clemens, Radelet and Bhavani (2004).

At the point where aid tends to zero the marginal product of aid will then be equal for all countries ( $k$ ) but its slope  $\beta_i$  will vary according to the measure of performance or capacity, of country  $i$ ,  $C_i$ .<sup>13</sup> If  $O$  is a measure of social outcomes and  $a$  is aid,

$$[1] \quad \frac{\partial O_i}{\partial a_i} = k - \beta_i a_i \quad [2] \quad \beta_i = \alpha(C_i)^{-\gamma}$$

Where  $\alpha > 0$  and  $\gamma > 1$  are scalars.

To take a specific example, we depart from the IDA allocation formula. At the margin, the cost of raising one dollar of IDA is one dollar; it is therefore reasonable to assume that at the optimal

<sup>13</sup> This implies that returns to aid can be negative where the quality of policies is low and aid is high. One mechanism could be through Dutch disease, if large aid inflows cause significant real exchange rate appreciation, which has high costs in terms of competitiveness, if aid is not used effectively. Another mechanism can be the effects of high aid flows on governance in weak institutional environments, in propping up poor policy regimes and unaccountable governments.

level and allocation of IDA, its marginal product will be one in all countries.<sup>14</sup> The optimal level of aid  $a_i^*$  is then given by [3], where  $C_i$  is country  $i$ 's CPIA score. Given a fixed stock of available aid,<sup>15</sup> the optimal aid allocation is fixed in proportion across countries according to the quality of policies and institutions, as [4].

$$[3] \quad a_i^* = (k-1)\alpha^{-1}(C_i)^\gamma \quad [4] \quad \frac{aid_1^*}{aid_2^*} = \left(\frac{C_1}{C_2}\right)^\gamma$$

The marginal product of aid falls to zero at some multiple  $m = k/(k-1)$  of the optimal aid level. This parameter,  $m$ , is central to the global debate over scaling up aid: how much can aid flows expand in the short-term while maintaining positive returns, given the current quality of policies and institutions? Clemens, Radelet and Bhavani's estimates (2004) suggest that the marginal product of aid falls to zero when aid is doubled from present levels: we term this the "optimistic" scenario. "Skeptics" who emphasize absorptive capacity constraints may believe that little productive increase in aid is immediately possible: for them,  $m$  is much closer to one. World Bank studies of absorptive capacity suggest that many countries are now able to productively absorb substantially more aid, say 50% above present levels. Without intending to imply a judgment, we term these the "realist" estimates.<sup>16</sup>

Figure 6 shows the marginal productivity of aid as  $k$  at  $a = 0$ , and falling at the rate  $\beta_i(C_i)$ . Optimal aid allocations  $a_1^*$ ,  $a_2^*$  and  $a_3^*$  correspond to  $C_1$ ,  $C_2$  and  $C_3$ . In high-performing countries ( $C_i = C_3$ ), the slope  $\beta$  is relatively flat, so that a large quantity of aid can be absorbed productively. In poorly governed countries ( $C_i = C_1$ ), the marginal product of aid drops rapidly. If actual aid allocations  $a_1$ ,  $a_2$  and  $a_3$  differ from optimal allocations, this causes efficiency losses of two types: countries may get too much aid, such that the marginal product is less than one (type 1 error), or too little, such that the marginal product is greater than one (type 2 error). These triangular losses are easily quantifiable:

$$[5] \quad L_i(a_i, a_i^*, m) = 0.5|1 - MPa_i||a_i - a_i^*|$$

This model provides a framework for thinking about the trade-offs between "optimality" and the stability of aid flows. Pre-commitment which guarantees a certain level of aid over several years is desirable in terms of stability and predictability; but it can generate efficiency losses if the quality of policies drifts over the horizon of the aid program. If country performance deteriorates sharply, aid will flow despite its falling social impact; similarly, if pre-commitments are upwardly rigid, countries which rapidly improve their performance may be denied funds that they otherwise might use well. Because rising aid levels confront absorptive capacity constraints more slowly in better-governed countries, both types of losses are smaller in such countries for misallocations of similar magnitudes.

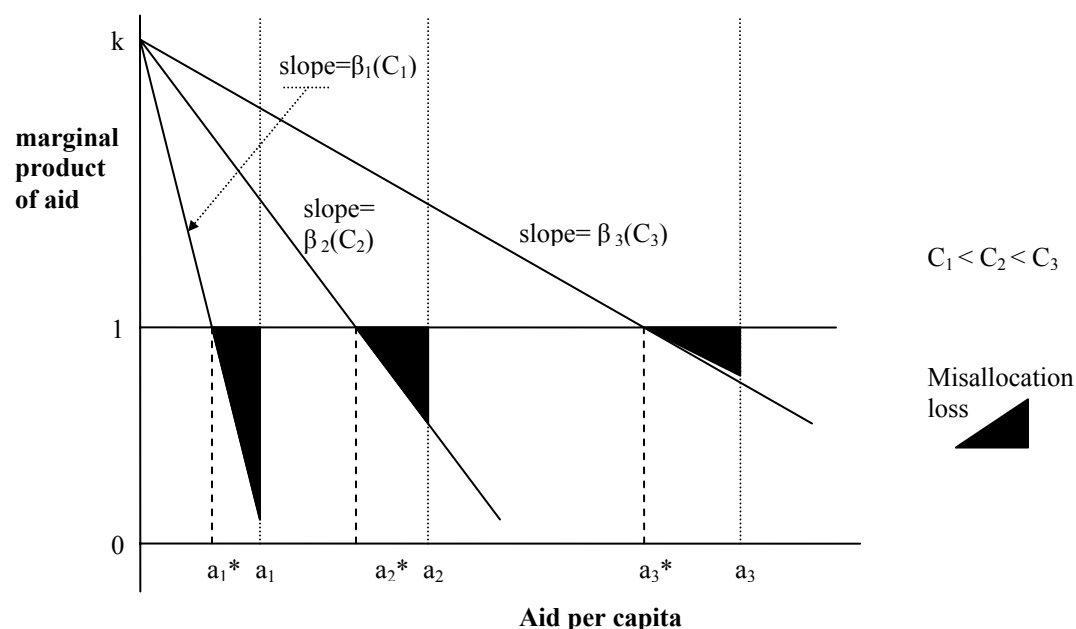
<sup>14</sup> This assumption could be relaxed, to set the marginal product of aid equal to the donors' shadow cost of aid, which may be higher or lower than one depending on the political pressures their governments face.

<sup>15</sup> Alternatively, this could be the globally optimal level of aid, if the marginal product indeed equals one.

<sup>16</sup> The parameter  $m$  should be distinguished from scaling-up proposals that allow some time for absorptive capacity and performance ratings to increase, so opening the way for further productive increases in aid.



**Figure 6 Aid Allocation and Efficiency**



*Simulating Trade-Offs in Aid Predictability*

We use five years of CPIA country performance ratings, for 1999-2003. Given this record, how much more stable would aid flows have been under a purely performance-oriented regime based on the model above, or under a modified regime with a flexible pre-commitment rule to improve predictability? How large would efficiency losses have been if five-year forward-looking aid programs had been enacted in 1999 according to various pre-commitment rules?

Tables 2a and 2b list countries considered in the exercise, breaking the group into CPIA quintiles for 1999 and for 2003. Roughly half of them remain in the same quintile, one-quarter move up, and one-quarter move down. Most movements are across a single quintile, but some countries slip more (for example, Cote d'Ivoire, Eritrea, Zimbabwe). Patterns of year-on-year changes in performance are important for the design of mechanisms to improve aid predictability. If one-year drops in CPIA scores tend to signal subsequent declines, the Type-I losses from not responding to initial signs of deterioration of performance could be substantial; similarly, large Type-II errors may occur if one-year improvements tend to foreshadow sustained improvements. On the other hand, if movements in the CPIA tend not to be auto-correlated (so that last year's change is not a good predictor of this year's change) then the case for not placing great stress on small year-on-year falls in CPIA scores may be strong. This is more typically the case; initial changes in ratings are often reversed and are not a good guide to future movements.

One reason for the random nature of short-term changes in performance rating can be measurement error. Using the natural experiment of two similar and independent performance assessments, one by the African development Bank and the other by the World Bank, Gelb, Ngo and Ye (2004) estimate the standard error of a CPIA estimate as about 2.4 points. While not

large in relation to the overall rating scale, this exceeds the typical annual change of plus or minus 0.1 point seen in the country sample. Most countries' annual changes are therefore within the range of measurement error, with a typical country requiring two, three or four years of sustained change to pull its rating outside this range. This has implications, discussed below, for the design of budget support.

Another reason for small movements can be changes in the survey instrument. Any assessment mechanism can be expected to evolve over time as knowledge accumulates, and even if the changes are gradual (as in the case of the CPIA) they can shift the relative position of countries slightly. This is an inevitable feature of a living assessment system as donors will want to take advantage of the best knowledge to allocate assistance.

We consider three values of  $m$ : the skeptic (1.25), the optimist (2), and the realist (1.5) values. Each doubling of the distance between  $m$  and 1 (that is, from 1.25 to 1.5, from 1.5 to 2) is associated with a halving of type-I and type-II losses. The "skeptic" will thus see efficiency losses twice as large as the "realist", and the "optimist" will see them half as large as the "realist".

The other parameter of interest is  $\gamma$ , the relationship between country performance (as proxied by the CPIA) and the slope of the marginal product function with respect to aid,  $\beta$ . The IDA allocation formula has a CPIA-aid elasticity of roughly 3; therefore, we use  $\gamma = 3$  throughout. We could also consider the case of direct budget support. There is no formula here, but because the better-performing countries tend to receive budget support, its allocation is more sensitive to performance than overall aid. We could take an approximate performance elasticity for budget support to be 4.

We can now fit the approximate IDA allocations in per head terms. Figure 7 shows the relationship between the CPIA, the slope of the marginal product of aid ( $\beta$ ), and the optimal level of aid in per capita terms ( $\text{aid}^*$ )<sup>17</sup>. The figure on the left is parameterized and scaled for overall IDA flows, and the figure on the right is closer to budget support. The mid-point of the CPIA scale, 3.5, is associated with overall aid flows of roughly \$10 per capita and budget support of \$4 per capita; both overall IDA and the share provided as budget support rise with the CPIA.

---

<sup>17</sup> These "normal" allocations abstract from special allocations to post-conflict countries. The approximation to the IDA rating system used also abstracts from the super-weighting of the governance-related component and the weighting on portfolio performance.

*Table 2a. Countries by CPIA Quintiles, 1999*

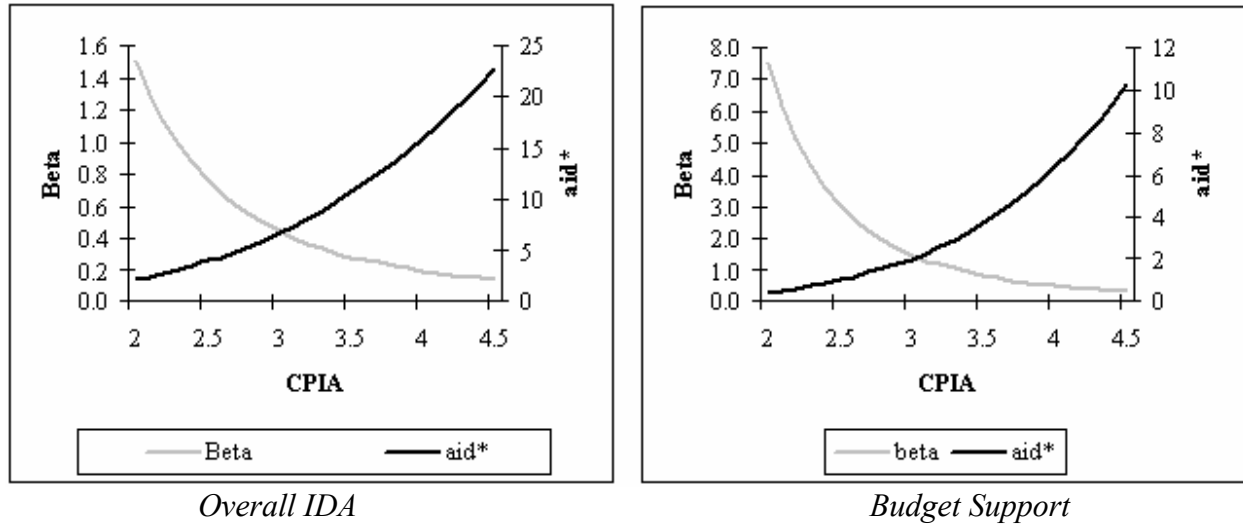
1	2	3	4	5
ARMENIA	BENIN	BURKINA FASO	AZERBAIJAN	ANGOLA
BANGLADESH	ERITREA	CAMEROON	CAMBODIA	BURUNDI
BHUTAN	ETHIOPIA	GAMBIA, THE	CHAD	CAR
BOLIVIA	INDIA	GEORGIA	DJIBOUTI	COMOROS
COTE D'IVOIRE	KYRGYZSTAN	KENYA	GUINEA	CONGO, DEM. REP.
ESTONIA	LESOTHO	MADAGASCAR	INDONESIA	CONGO, REP.
GHANA	MALAWI	MALI	KIRIBATI	EQUATORIAL GUIN.
GUYANA	MOZAMBIQUE	MOLDOVA	LAO, PDR	GUINEA-BISSAU
HONDURAS	NICARAGUA	MONGOLIA	NEPAL	HAITI
MAURITANIA	TANZANIA	PAKISTAN	NIGER	PAPUA NEW GUINEA
SENEGAL	YEMEN	RWANDA	NIGERIA	SAO TOME & PR.
SRI LANKA	ZAMBIA	SOLOMON ISLA	TOGO	SIERRA LEONE
UGANDA		VIETNAM	ZIMBABWE	SUDAN
				TAJIKISTAN
				UKRAINE
				UZBEKISTAN

*Table 2b. Countries by CPIA Quintiles, 2003*

1	2	3	4	5
ARMENIA	<b>AZERBAIJAN</b>	CAMEROON	CAMBODIA	ANGOLA
BHUTAN	<b>BANGLADESH</b>	<b>ETHIOPIA</b>	CHAD	BURUNDI
ESTONIA	BENIN	GEORGIA	<b>CONGO, DR</b>	CAR
HONDURAS	<b>BOLIVIA</b>	<b>GUYANA</b>	<b>CONGO, REP.</b>	COMOROS
<b>INDIA</b>	<b>BURKINA FASO</b>	KENYA	<b>COTE D'IVOIRE</b>	EQUATORIAL GUIN.
MAURITANIA	<b>GHANA</b>	<b>KYRGYZSTAN</b>	DJIBOUTI	GUINEA-BISSAU
<b>NICARAGUA</b>	<b>INDONESIA</b>	<b>LESOTHO</b>	<b>ERITREA</b>	HAITI
SENEGAL	<b>MADAGASCAR</b>	<b>MALAWI</b>	<b>GAMBIA, THE</b>	<b>LAO, PDR</b>
SRI LANKA	<b>MALI</b>	MOLDOVA	GUINEA	<b>NIGERIA</b>
<b>TANZANIA</b>	<b>NEPAL</b>	MONGOLIA	KIRIBATI	PAPUA NEW GUINEA
UGANDA	<b>PAKISTAN</b>	<b>MOZAMBIQUE</b>	NIGER	SAO TOME AND PR.
<b>UKRAINE</b>	YEMEN	RWANDA	<b>SIERRA LEONE</b>	<b>SOLOMON ISLA</b>
<b>VIETNAM</b>		<b>ZAMBIA</b>	<b>TAJIKISTAN</b>	SUDAN
				<b>TOGO</b>
				UZBEKISTAN
				<b>ZIMBABWE</b>

Quintile unchanged; **Quintile improved**; *Quintile worsened*.

**Figure 7. Optimal Aid Allocations and the Quality of Policies and Institutions**



#### *Simulating Losses from Pre-commitment*

The CPIA has been broadly consistent over periods of time, but has changed slightly from year to year. There is some slight upward drift in the 1999-2003 ratings and we normalize the scores for each year so that the means are consistent across years.<sup>18</sup> We assume that aid in 1999 reflects the optimal allocation (from equation [3] above) and consider three types of programs for 2000-03. The first, the annual performance-based system, sets  $aid_{it} = aid_{it}^*$  for each country and year. The second, the pure pre-commitment system, holds each country's aid level 2000-2003 at its (optimal) 1999 level. The third, the flexible pre-commitment rule, adjusts aid levels if and only if a country's CPIA score drifts to a point that is well above or below its 1999 level. Here, this critical point corresponds to  $\pm 0.33$ , around a 92% confidence interval associated with the estimated standard error of the CPIA.<sup>19</sup> This rule will not adjust aid flows unless a country has had a rather clearly observable performance improvement or a major performance deterioration. For each of these three programs, we perform two sets of analyses.

First, we compute the degree of volatility inherent in the "optimal" performance-based aid flow. If we had a purely performance-based system without additional safeguards for predictability, how stable would aid flows be? How does this level of stability compare to the stability of aid in recent years? How much is stability further improved by the flexible pre-commitment rule? Second, we compute the implied efficiency losses from aid misallocations over 2000-03 arising from the pure and flexible pre-commitment rules given countries' actual CPIA trajectories.

<sup>18</sup> Normalization is needed to eliminate biases from changes in overall performance trends over time. To some degree these changes are real, but they can also result from slow evolution in the rating instrument. Some degree of evolution is inevitable in any rating scale, and is considered part of a normal process of re-assessment.

<sup>19</sup> The one-tail significance level is appropriate here.

### *Stability of Aid Flows*

Performance-based aid should be volatile if country performance is volatile. Using the model laid out above, we simulate the volatility of an optimal performance-based system with no stability safeguards, and compare it to that of the system with the flexible pre-commitment rule. Of course, this question is irrelevant for the full pre-commitment rule, which eliminates all variability.

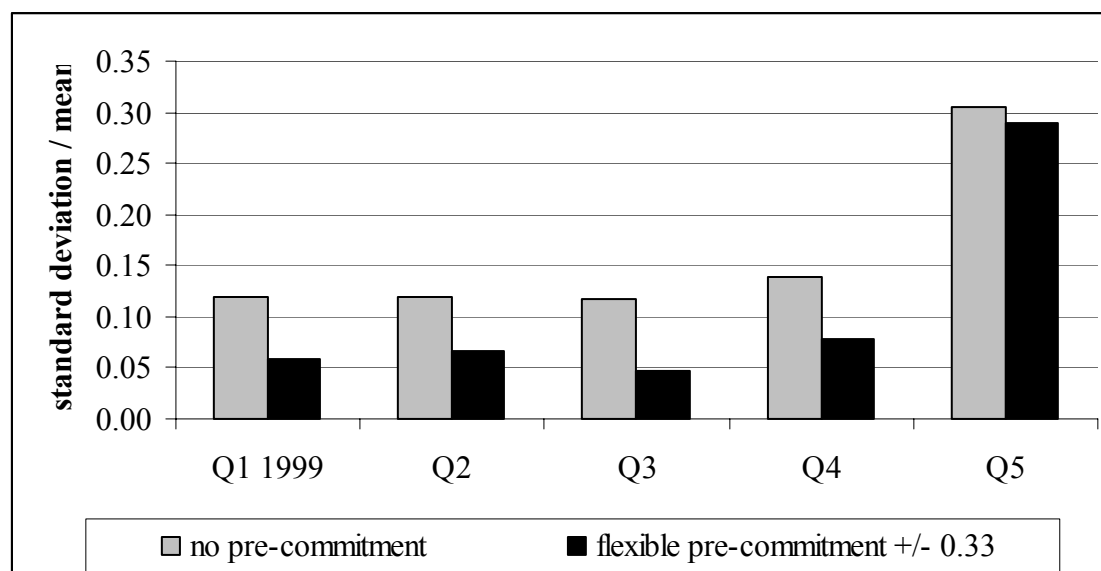
Under the optimal allocation system with no pre-commitments, aid allocations have an average standard deviation of 17 percent of 1999 levels. Much of this volatility comes from the lowest-performing countries; for the first four 1999 CPIA quintiles, the standard deviation is between 0.12-0.14, compared to 0.31 for the last quintile (Figure 8). Being in the initial CPIA quintile does not make much of a difference for simulated volatility; over the last five years, a number of countries in the top two quintiles have slipped badly (Table 2). Countries which perform consistently – whether first quintile or third – benefit from low volatility. In countries with initially worse performance, subsequent improvements result in rapid increases in aid flows and hence greater volatility.

In comparison, the estimates of historical volatility for most low-income countries from Bulir and Hamann (2003) are in the range of 0.40 – 0.60, or 0.30 – 0.40 around an autoregressive forecast and they suggest that these trends extend through recent years as well. This implies that a purely performance-based system, even without pre-commitments, would be a vast improvement in terms of predictability over the current aid regime: reductions in volatility for most countries would be on the order of two thirds of their past levels.

As Figure 8 also shows, the flexible pre-commitment rule is quite successful in reducing volatility where performance deviations are modest. Our simulations suggest that it is capable of delivering another 50 percent reduction in average variability for countries in the top four quintiles. For those countries which stay on-track throughout the program, it reduces variability all the way to zero. Of course, where programs go rapidly off-track (as in several of the quintile-five countries), the flexible rule has little effect. One reason for not providing budget support, or much budget support, to low-performing countries is that it is harder to reconcile stable financing with performance-based allocation.

In summary, a performance-based allocation mechanism based on the model developed above would have substantially improved the predictability of existing aid allocation mechanisms, reducing the variability of flows to a third of their historical level. The addition of a flexible pre-commitment rule, in which five-year forward-looking aid programs are guaranteed on the basis of current performance as long as performance does not vary outside of the 92% confidence interval, would have further reduced the variability of aid flows by half.

**Figure 8. Stability of Aid Flows, Standard Deviation 1999-2003, share of 1999 aid flow**



### *The Efficiency Trade-Off*

We now impose pure and flexible pre-commitment rules on the 2000-03 programs, and compute the efficiency losses from aid misallocations over 2000-03 arising from the rules' responses to divergence between original and subsequent performance.<sup>20</sup> Tables 3a and 3b present the results for the pure pre-commitment rule, where  $a_t = a_{1999}(C_{1999})$  for all  $t$ . For simplicity, the discussion will use the "realist" assumption that  $m = 1.5$ ; "skeptics" can double the estimated losses and "optimists" can halve them.

More than half of Type-I losses are moderate, equivalent to less than 5% of the total aid flows to the recipient country. Better-performing countries in 1999 are not necessarily less likely to incur moderate losses than their worse-performing counterparts, in part because their potential mobility is mostly downward. In fact, substantial Type-I errors can occur where strong performers suddenly deteriorate: Cote d'Ivoire's slip into conflict pushed it from the first quintile to the fourth, while Guyana slid from the first quintile to the third. Pre-commitments in these countries would have engendered large efficiency losses. However, large Type-I efficiency losses are rare, and occur only in countries which were initially not too highly rated but slip further (Solomon Islands, Zimbabwe). The total sum of Type-I losses generated by the pure pre-commitment program, weighted by the amount of aid received in 1999, is about 4.2% of total aid flows, with no strong pattern by initial CPIA score (Table 4).

<sup>20</sup> A further concern may be the set of arguments related to the influence of aid on the quality of policies and institutions. On one hand, aid in the form of budget support and public sector management projects might plausibly be used to support programs to improve government capacity. On the other hand, aid which does not respond adequately to government performance might plausibly have a negative influence on the quality of policies by softening the government budget constraint. We return in part to this relationship in Section 6.

Type-II errors are less common than Type-I errors, but often larger in magnitude. This is because the average magnitude of improvements, though rarer, was larger than the average magnitude of slippage. Large Type-II errors never occur in well-managed countries, first because of the flatness of their marginal productivity curves ( $\beta_i$ 's), and second because they have little upward mobility on the CPIA. In contrast, type-II errors can be very large in initially poor-performing countries which make progress over the horizon of the four-year program. Almost half of type-II errors occur in countries which started in the fifth CPIA quintile in 1999, and several are equivalent to more than 25% of aid flows, including DRC, Republic of Congo and Ukraine. The total sum of Type-II losses generated by the pure pre-commitment program, weighted by the amount of aid received in 1999, is about 6.5% of total aid flows, almost entirely incurred in countries with low initial CPIA scores (Table 4).

The sum of Type I and Type II errors under full commitment is therefore 10.7 percent of aid flows, a not-insignificant number.

**Table 3a. Type-I Losses under Pure Pre-Commitment**

Loss, % of aid	Q1 1999	Q2 1999	Q3 1999	Q4 1999	Q5 1999
0-2		India	Gambia		
2-5	Bolivia	Lesotho, Malawi, Mozambique, Nicaragua, Zambia		Togo	
5-10	Bangladesh, Ghana		Cameroon	Lao PDR, Nigeria	
10-25	Cote d'Ivoire, Guyana	Eritrea, Ethiopia, Kyrgyzstan			
25-100					
>100			Solomon Islands	Zimbabwe	

**Table 3b. Type-II Losses under Pure Pre-Commitment**

Loss, % of aid	Q1 1999	Q2 1999	Q3 1999	Q4 1999	Q5 1999
0-2		Tanzania	Burkina Faso, Madagascar, Pakistan		
2-5			Vietnam	Indonesia	
5-10					Tajikistan
10-25				Nepal	Sierra Leone
25-100					
>100					Congo DR, Congo R, Ukraine

**Table 4. Simulated efficiency losses, % of aid flows for 2000-2003, pure pre-commitment**

Quintile	Type-I			Type-II		
	$m = 1.25$	1.5	2	$m = 1.25$	1.5	2
1	9.2	4.6	2.3	0.3	0.1	0.1
2	7.4	3.7	1.8	0.1	0.1	0.0
3	3.5	1.8	0.9	2.4	1.2	0.6
4	12.9	6.5	3.2	7.0	3.5	1.7
5	10.1	5.0	2.5	107.7	53.8	26.9
all	8.3	4.2	2.1	13.0	6.5	3.8

Tables 5a and 5b present the results for the flexible pre-commitment rule. This rule cuts efficiency losses dramatically. No country generates losses equivalent to more than 10% of its aid flow, and most lie in the range of 0-5%. Countries which experienced rapid deterioration in governance, and thus generated large losses under the pure pre-commitment rule, now go off-track for most or all of the program; the flexible aid system also scales up assistance to countries which see major improvements in performance from a very low level. The sum of Type-I losses falls to 1.3% of total aid flows, or less than a third of its level under pure pre-commitment (Table 6). Type-II losses are very small (0-2% of aid flows) and concentrated in the bottom three quintiles. Total losses using the flexible pre-commitment rule are only 2.3 percent of aid flows.

**Table 5a. Type-I losses under flexible pre-commitment**

Loss, % of aid	Q1 1999	Q2 1999	Q3 1999	Q4 1999	Q5 1999
0-2		India, Kyrgyzstan	Gambia		Sierra Leone
2-5		Eritrea, Lesotho, Malawi, Mozambique, Nicaragua, Zambia	Vietnam	Togo	Tajikistan
5-10	Bangladesh, Ghana			Lao PDR, Nigeria	

**Table 5b. Type-II losses under flexible pre-commitment**

Loss, % of aid	Q1 1999	Q2 1999	Q3 1999	Q4 1999	Q5 1999
0-2		Tanzania	Burkina Faso, Madagascar, Mali, Pakistan	Nepal	Congo DR
2-5	Bolivia			Azerbaijan, Indonesia	

**Table 6. Simulated efficiency losses, % of aid flows for 2000-2003, flexible pre-commitment**

Quintile	Type-I			Type-II		
	$m = 1.25$	1.5	2	1.25	1.5	2
1	5.4	2.7	1.3	0.2	0.1	0.05
2	3.8	1.9	0.9	0.2	0.1	0.05
3	0.6	0.3	0.2	2.5	1.2	0.6
4	1.6	0.8	0.4	4.4	2.2	1.1
5	1.0	0.5	0.2	2.8	1.4	0.7
all	2.7	1.3	0.7	2.0	1.0	0.5



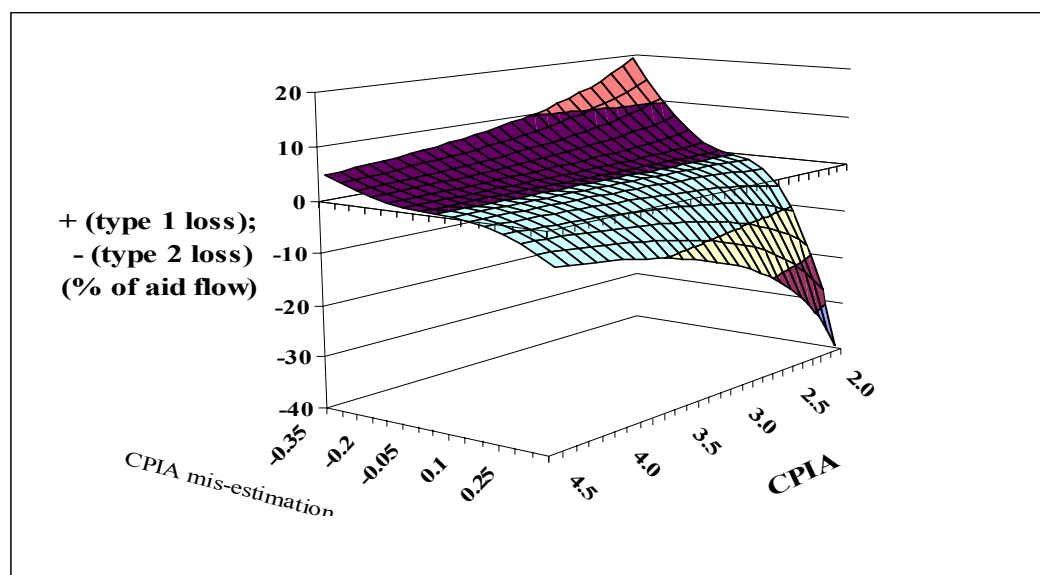
*Measurement Error in the CPIA Process: Implications for Efficiency*

As noted above, even if the CPIA is the “right” measure of country performance and is estimated through a rigorous process of judgment, along with any other performance estimate (whether policy-based or output/outcome-based) it will be subject to measurement error. This nuances the task of aid delivery significantly. It creates the potential for mis-allocation even without pre-commitment rules, because ratings may be too high or too low; It also obscures the judgment of whether country performance has improved or worsened.

How large are the potential efficiency losses from mis-allocations of aid which arise from errors in the measurement of the CPIA? This is a simple exercise: for each level of the “true” or underlying CPIA, we consider the efficiency costs of providing aid at the inaccurate levels around the “true” CPIA (within the 92% confidence interval). Figure 4.3 shows the results, with  $\gamma = 3$  and  $m = 1.5$ . The inverse relationship between performance levels and losses from mis-allocation is evident -- even at the maximum mis-judgment ( $\pm 0.33$ ), losses are only 5-6% of aid flows for countries in the high CPIA range (4 – 4.5), compared to 10-20% in the low range (2 – 3), and even higher on the Type-II side for very weak performers.

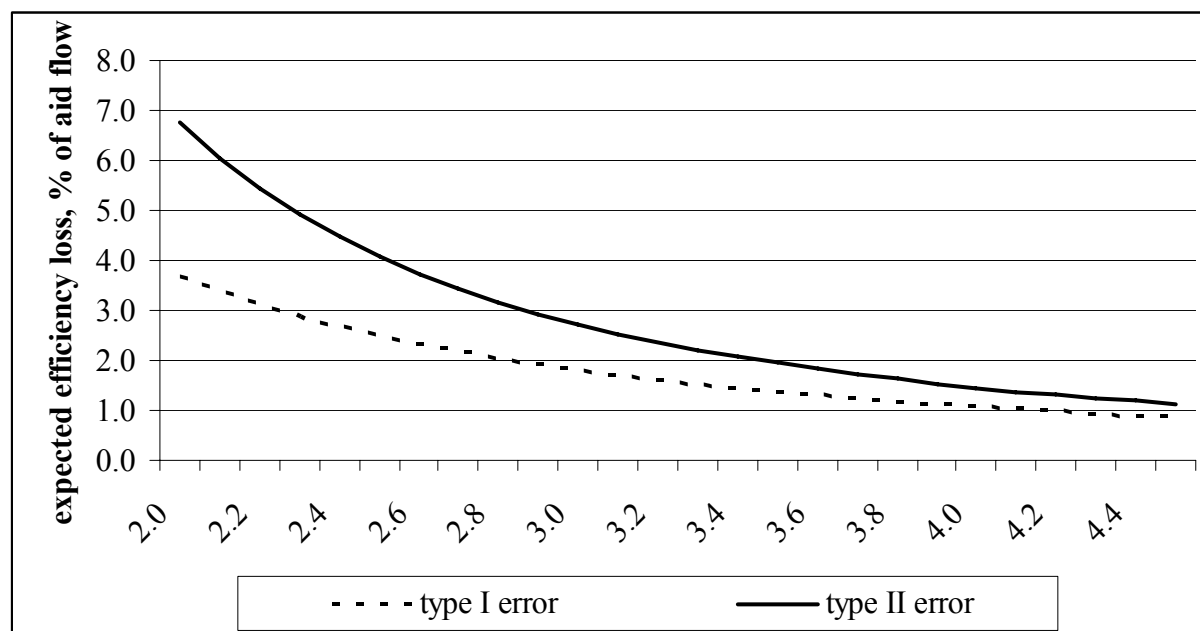
Figure 10 shows the expected value of Type-I and Type-II efficiency losses stemming from mis-estimation of the CPIA, assuming that errors follow a normal distribution. For the poorest-performing countries these add up to more than 10% of the value of the total aid flow. For the bulk of the countries, clustered in the range of 3.25 – 3.75, expected efficiency losses are roughly 3-4% of the aid flow.

**Figure 9. Potential efficiency losses in optimal allocation system from errors in CPIA**



\*  $m = 1.5$

**Figure 10. Expected value of efficiency losses from mis-estimation, by CPIA level**



The likely efficiency losses from mis-measurement are therefore larger than the previous estimates of efficiency losses from misallocation under the flexible pre-commitment rule. This result suggests that, given the imprecise measurement of performance, there is little to gain from continually fine-tuning aid levels in response to minor year-on-year performance changes. On the other hand, such a graduated response needs to include a mechanism to reassess the level of support when performance clearly improves or deteriorates by a large margin.

## 5 Calibrating Budget Support -- Performance Levels versus Changes?

Given a system for allocating overall aid, how should the fraction going through budget support, rather than project support, be determined? Country circumstances will partly shape the answer but some common principles may apply. Most donors, including the Bank, tend to channel more aid to better-performing countries through budget support. A recent review of Poverty Reduction Support Credits (PRSCs) shows quite a high degree of selectivity, with most recipients being in the top three quintiles of the CPIA ratings and only one (Guyana) in the third quintile.<sup>21</sup> Budget support may thus amount to up to half of total aid for high-performing countries but to little or nothing in the low performers.

Such selectivity in the allocation of budget support can be implemented in several ways. One approach would be formula-based, where countries become eligible above a certain performance cut-off, and where the maximum share increases with levels of performance. With overall IDA allocations proportional to the third power of the CPIA (as in the previous section), setting

<sup>21</sup> World Bank (2005), See also the progression of support instruments set out in World Bank (2004).

budget support proportional to the fourth power of the CPIA would approximate IDA's current practice. Assuming budget support is valued highly by recipients relative to project support, this would also provide an additional incentive for countries to move up the performance rankings. The costs of pre-commitment using a fourth-power allocation system can also be considered using the model of the previous section; they too are modest, especially for the higher-rated countries, and using the flexible pre-commitment rule that responds to only major performance changes. Budget support, too, could be committed based on performance levels, with a provision that enables a reassessment in response to major changes.

However, such an approach may not fully capture the important role of budget support as an investment in country systems. In some cases, donors agree to channel funds through still-fragile country systems with the expectation that countries will use this opportunity to strengthen their own systems of budget and financial management and service delivery. At very low levels of capacity, the expected cost in terms of ineffective or corrupt use of funds exceeds the likely benefits (and in addition it will be more difficult to reconcile performance-based assistance with stable financing). But as country capacity increases, donors will find it appropriate to encourage capacity improvements through budget support in countries on a promising trajectory even if they are not comfortable with the existing levels of budget and financial management and service delivery. The more rapid are the capacity gains the greater is the return on the donors' willingness to take risks. Without demonstrated improvements in capacity that are high enough to yield an acceptable return on their financial and reputational investment, donors can credibly withdraw from budget support back towards projects, but only if the level of capacity is still below their comfort threshold. For very high-capacity clients, donors can provide assistance through budget support simply by "certifying" country systems.<sup>22</sup> Unless facing a large deterioration in performance, it is not credible for them to threaten to return to projects in such countries.

This notion of budget support as an investment in country systems suggests that criteria for budget support should reflect both **levels** and **trends** in performance. This stands generally in contrast to the CPIA and similar systems, including the indicator-based MCA criteria, which focus exclusively on levels, though with two exceptions. First, levels of performance tend to be interpreted in a comparative context, so a country that lags all others in improving performance will see its allocation fall even if there is no absolute change in its performance. Second, depending on progress against CAS triggers, countries can move between low, base and high cases in the IDA system, with variations of up to about 30% of the base case over a 3-4 year CAS period; countries can also see commitments frontloaded or backloaded in response to anticipated changes in performance and allocations. Such changes can be interpreted as compensating, to a degree, for the backward-looking nature of the CPIA process. This approach can be consistent with seeing budget support as an investment in country systems in response to positive trends.

This story is illustrated in Figure 11. With assistance delivered through traditional project mechanisms, capacity improves slowly along the line ABC. Only at time T2 does it reach the level C\* at which donors feel fully comfortable in channeling assistance through the budget. At

---

<sup>22</sup> This approach has been suggested for IBRD support to some middle income countries with high quality policies and institutions.

some earlier time, T1, however, the country's track record is strong enough so donors can begin to shift to budget support, hopefully creating a "virtuous circle" that will further strengthen country capacity, as along ABDC. The risks to the donors of providing support during the time when capacity lies below their level of comfort, are offset by the potential gains from the more rapid increase in capacity levels to above their comfort threshold. The slope of BD determines the return to the donors' investment, so that the speed of improvements in country systems (not just the level) is an important indicator of effectiveness for budget support. Of course, the slope of BD is unknown *a priori*; the previous track record and evidence of government commitment will be central to donors' decision to take the risk of providing budget support.

The question then arises of the relative weighting of performance levels and trends. Too small a response to positive trends might not reinforce the need to improve systems and coordinate ministries to implement agreed programs: such reinforcement is welcomed by some clients.<sup>23</sup> But too strong a response reduces predictability, thus undermining the value of the budget support instrument and eliminating its potential incentive effects. There is no simple answer, especially given that it is not always easy to distinguish trends in performance from temporary changes or measurement errors. But the analysis above of the likely size of error in measuring performance and the benefits and costs of flexible pre-commitment suggest ways to approach the problem.

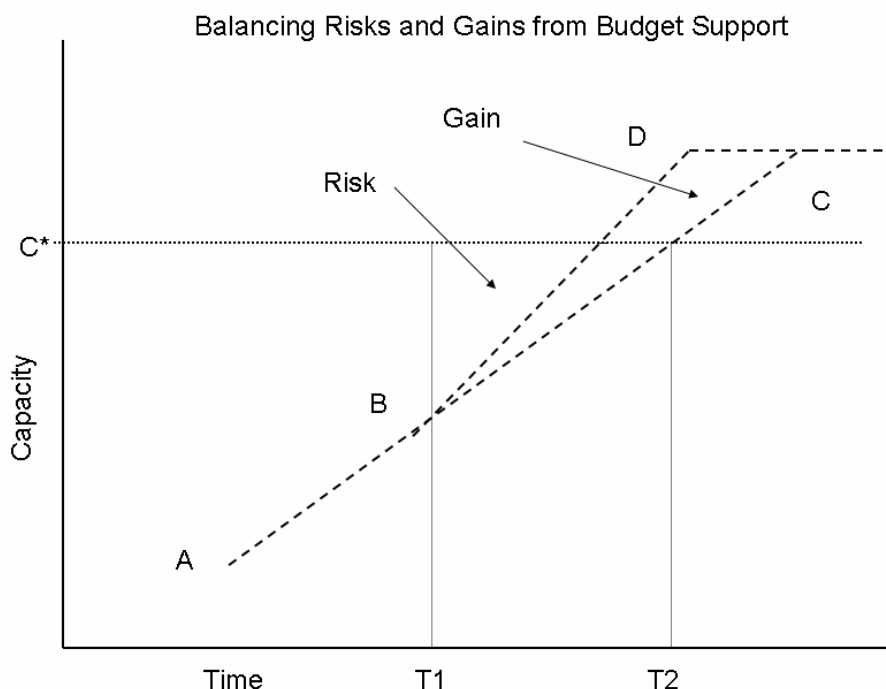
One approach would be for donors to establish a code of good practice. This would set a base level of budget support for several years at a time and supplement it by incentive payments based on agreed actions or other interim assessments of performance. These payments would be modest, perhaps up to 10 percent of base support, corresponding to the modest changes in overall assistance that would follow "normal" (if uncertain) annual changes in a CPIA. They would thus anticipate future changes in the envelope, and be applied to the next year's support rather than the current level in order to further improve the predictability of assistance.

Every three years or so, there would be a deep, systematic review of progress in strengthening country systems of budget and financial management and service delivery, support by independent assessment and comprehensive output measurement. This would feed back into the CPIA, and help shape the decision on how much future support to channel through budget support. As in the case of overall aid, major performance changes, whether positive or negative, should trigger a comprehensive review, informed by neutral external assessment. The results of Section 4 suggest that flexibly committing support forward in this way will involve little efficiency loss while greatly increasing predictability.<sup>24</sup>

---

<sup>23</sup> Surveys conducted by the SPA Working Group on Budget Support suggest that recipients are not necessarily opposed to conditionality. For instance, conditions can bolster the positions of ministers to pursue needed reforms, and create incentives to work across ministerial lines. On a 1-5 scale, the usefulness of conditionality was rated as 3.1 in 2002/3 and as 3.6 in 2003/4. The intrusiveness and specificity of some conditions may cause concern, so can an excessive number, and the instability of financing that can result from high leverage on conditionality. The last is an issue of instrument design however, not one of conditionality per se.

<sup>24</sup> This approach could be implemented in a number of possible ways. One would be to separate budget support into a set of well-defined segments, each responding to performance assessments for a given sector (World Bank 2004). The difficulty of leveraging this strongly on an annual basis is that one year provides a short interval for definitively assessing efficiency changes.

**Figure 11**

## 6 Developing Performance-Based Norms to Guide Aid Allocations

A final issue concerns the distinction between policy-based and results-based aid allocations. Within the development community, recent years have seen a shift away from emphasizing only policy actions and towards including a focus on results. The European Community's budget support programs represent the most ambitious move in this direction: they combine a fixed tranche, which is delivered as long as a set of basic conditions are met, with a variable tranche which disburses at a level determined by the recipient country's success in meeting a set of mutually-agreed targets for service delivery and public financial management. As shown in the recent assessment of experience (European Commission 2005), the approach has been quite successful in combining a reasonable degree of predictability with performance-based incentives.

At the core of outcome-based approaches is the set of targets which determine the disbursement of performance-related assistance. As emphasized by EC's evaluation of its programs, there is as yet no analytic framework guiding the setting of targets. What is an appropriate three-year target for improvements in the primary enrolment rate, given the current enrolment rate and possibly other variables like income? How rapidly can schooling quality improve, or child mortality decline? There has been little research on the pace of change in outcome, output and service-delivery indicators. The EC evaluation identifies the issue of target-setting as one of the most important remaining issues for its new framework.

This section considers how targets might be set with reference to historical evidence. EC programs use indicators for direct government outputs (school enrollment rates, number of children immunized) rather than broader impacts like infant mortality, but we take the examples of primary enrollment and under-five and infant mortality because of current data availability.<sup>25</sup>

### *Primary enrollment*

Clemens (2004) assembles evidence on the speed with which countries have moved from low to high levels of primary enrollment, fitting rates of improvement to a logistic function. He finds a strong pattern of regularity in their transitions, with a typical country taking 115 years to move from enrollment rates of around 10 percent to 90 percent. Without considering the influence of income-related and other factors, the normal speed of percent increase in enrollment is:

$$N = a(s)(1-s)$$

Where  $s$  is the current enrollment share. The parameter  $a$  is estimated at 3.8, with standard deviation 0.33. This means that, for a country starting off with 50 percent enrollment, “normal” progression is at 0.95 percentage points annually, with a 95 percent confidence interval of [1.12 percent, 0.78 percent]. This can be further refined by allowing  $a$  to vary between countries according to exogenous characteristics and policy-related factors. Clemens finds some sensitivity to income levels but little response to indicators of education policy. He further suggests that observed episodes of extremely high enrollment expansion relative to the norm are likely to reflect wrong or misleading data (for example, in one example, enrollment rates rose rapidly because large numbers of children were not allowed to exit from the system) or to be associated with sharp deterioration in indicators of educational quality.

Clemens’ norms do not show that rapid progress is impossible, but they can provide a useful reality check on the goals embedded in country programs. The targets embedded in the EU’s budget support programs for Burkina and Ethiopia, for example, are set at about 350% of the Clemens norms for these countries. Recent trends suggest that, while Burkina’s “slow” performance has been quite close to its norm, Ethiopia has expanded enrollment far more rapidly than the Clemens norm. The past might not be a good guide to the future, but experience suggests that sustaining progress at the rates targeted will be very challenging, especially if quality is to be sustained.

### *Infant and Under-5 mortality*

In this exercise, we compare mortality rates (infant and under-5, respectively) with subsequent (annualized) changes in mortality rates. Observations for mortality rates are for all countries for

---

<sup>25</sup> Infant mortality is not strictly speaking a service delivery indicator and is therefore not used in the budget support operations of the EU. However it is easily and consistently available across countries and over time, and is therefore used as an additional illustration. With more data effort, norms could be derived for a number of service delivery indicators, including immunizations, performance on standardized tests, etc. For an assessment of the linkages between governance, spending, growth and some of the MDG variables, see Baldacci, Clemens, Gupta and Cui (2004).

the years 1960, 1970, 1980, 1990, 1995 and 2000; the subsequent changes are measured as the average annual rate of change between those one-year observations. Clearly, where mortality rates are already low, subsequent changes tend to be small, but at higher mortality rate levels we see the subsequent rates of change diverging between countries. In poorer countries with high mortality rates, improved living standards or strong efforts to improve health outcomes can cause declines in mortality rates in the range of 3-5 per 1000 per year. In other countries, rates stagnate or even rise further, generally because high mortality rates are a symptom of persistent problems (conflict, poor governance) and/or because new challenges are emerging (AIDS).

Quantile regressions examining the relationship between current mortality levels and subsequent changes in mortality rates were executed at the 10<sup>th</sup>, 25<sup>th</sup>, 50<sup>th</sup>, 75<sup>th</sup> and 90<sup>th</sup> percentiles.<sup>26</sup> Tables 7a and 7b report the results. The coefficients  $\alpha$  are high at the 90<sup>th</sup> percentile (-1.1 and -1.7), and fall with lower percentiles, until the 10<sup>th</sup> percentile where there is no relationship. Coefficients are all statistically significant at the 99.9% level, except of course the 10<sup>th</sup>.

**Table 7a. Quantile Regressions on Annual Rate of Change in Infant Mortality**

percentile →	90		75		50		25		10
	$\alpha$	T	$\alpha$	T	$\alpha$	T	$\alpha$	T	$\alpha$
log infant mortality	-1.07	18	-0.8	22	-0.58	24	-0.29	11.6	0
constant	1.5	6.6	1.33	8.9	0.99	10.4	0.50	5.1	0
r <sup>2</sup>	0.28		0.28		0.26		0.06		0

**Table 7b. Quantile Regressions on Annual Rate of Change in Under-5 Mortality**

percentile →	90		75		50		25		10
	$\alpha$	T	$\alpha$	T	$\alpha$	T	$\alpha$	T	$\alpha$
log infant mortality	-1.7	24	-1.27	23	-0.88	30	-0.36	18	0
constant	2.84	10	2.3	10	1.77	15	0.65	8	0
r <sup>2</sup>	0.34		0.31		0.22		0.06		0

Table 8 illustrates the implications for improvements in mortality rates. Countries with relatively high mortality rates, if performing well, can make rapid progress on both infant and under-5 mortality. These estimates can be used to suggest appropriate target rates of improvement for forward-looking programs. On infant mortality, countries with mortality rates above 100 (as per most African countries) can reduce mortality at a rate of 2.5 – 3.5 per 1000 per year with strong efforts (75<sup>th</sup> percentile) and even faster (3.5 – 4.5 per 1000 per year) at the 90<sup>th</sup> percentile. These might be suitable goals for countries with sufficient capacity to receive budget support. For countries with particularly weak systems and difficult circumstances, the median may be more appropriate, where infant mortality rates improve by 1.7 – 2.3 per 1000 per year.

<sup>26</sup> See Koenker and Basset (1978)

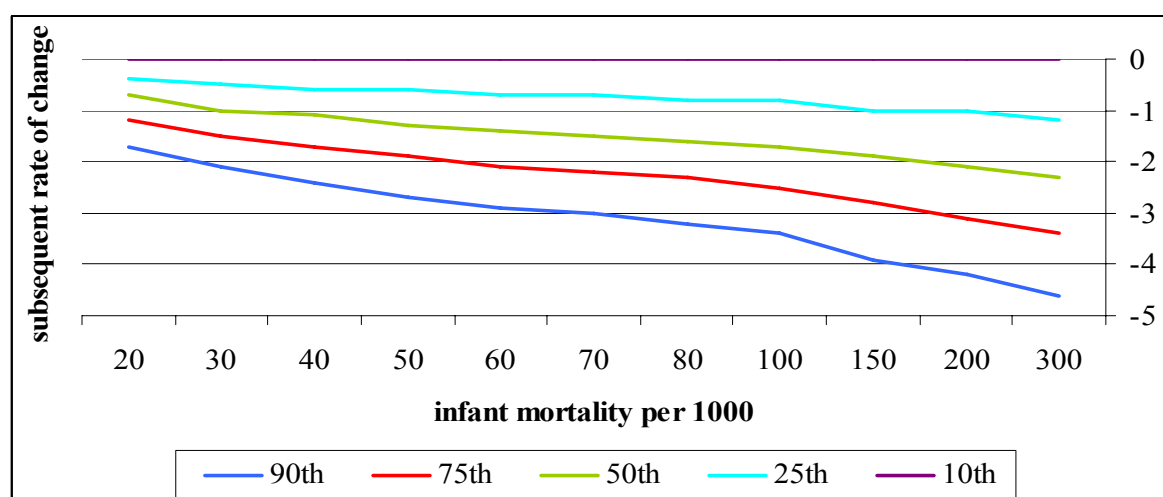
Patterns in the estimates for improvement on under-5 mortality are similar. The 90<sup>th</sup> percentile of countries in the 100+ range historically have seen improvements of 5 – 7 per 1000 per year in under-5 mortality, 3.5 – 5 at the 75<sup>th</sup> percentile, and 2.3 – 3.2 at the median.

Figures 12a and 12b illustrate the estimated “paths” from high to low mortality rates at different percentiles. The 90<sup>th</sup> percentile paths, for instance, suggest that a very strong performer starting with an infant mortality rate of 150 and an under-5 mortality rate of 250 could reduce those mortality rates to 80 and 135 respectively in 20 years. Over the same period, the 75<sup>th</sup> percentile paths would bring a country to 100 and 165 respectively.

**Table 8. Annual Rates of Improvement in Infant Mortality (per 1000)**

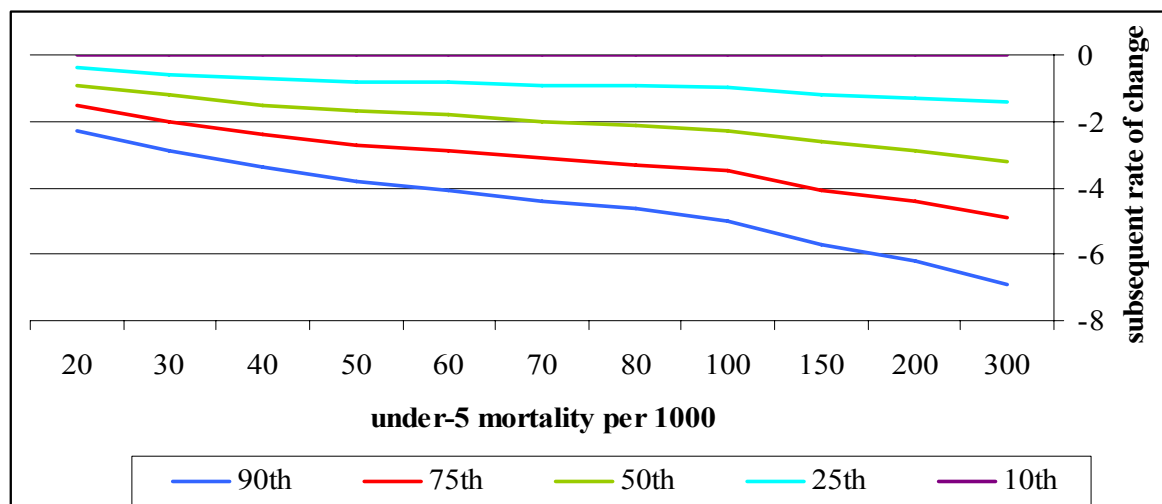
Initial mortality rate (per 1000)	Subsequent annualized change in infant mortality, at percentile...					Subsequent annualized change in under-5 mortality, at percentile...				
	90 <sup>th</sup>	75 <sup>th</sup>	50 <sup>th</sup>	25 <sup>th</sup>	10 <sup>th</sup>	90 <sup>th</sup>	75 <sup>th</sup>	50 <sup>th</sup>	25 <sup>th</sup>	10 <sup>th</sup>
300	4.6	3.4	2.3	1.2	0	6.9	4.9	3.2	1.4	0
200	4.2	3.1	2.1	1.0	0	6.2	4.4	2.9	1.3	0
150	3.9	2.8	1.9	1.0	0	5.7	4.1	2.6	1.2	0
100	3.4	2.5	1.7	0.8	0	5.0	3.5	2.3	1.0	0
80	3.2	2.3	1.6	0.8	0	4.6	3.3	2.1	0.9	0
70	3.0	2.2	1.5	0.7	0	4.4	3.1	2.0	0.9	0
60	2.9	2.1	1.4	0.7	0	4.1	2.9	1.8	0.8	0
50	2.7	1.9	1.3	0.6	0	3.8	2.7	1.7	0.8	0
40	2.4	1.7	1.1	0.6	0	3.4	2.4	1.5	0.7	0
30	2.1	1.5	1.0	0.5	0	2.9	2.0	1.2	0.6	0
20	1.7	1.2	0.7	0.4	0	2.3	1.5	0.9	0.4	0

**Figure 12a. Infant mortality rates, estimated paths at varying performance quantiles**





*Figure 12b. Under-5 mortality rates, estimated paths at varying performance quantiles*



## 7 Conclusion

This paper has considered three problems that will remain even if donors are able to lengthen their funding horizons to create multi-year aid budgets:

- First, how to deal better with exogenous factors that cause disbursements to diverge from commitments in the short-medium term?
- Second, how to commit aid forward in a multi-year framework that balances out the need for predictable funding against the risk of misallocating aid as countries' performance changes?
- Third, within the overall aid envelope, how to shape the breakdown of assistance into project support and budget support, in a way that responds to performance yet provides the degree of funding predictability needed for budget support to be effective?

On the first topic, we show that much could be done using active reserve management and fiscal rules to cushion spending against fluctuations in disbursements that are not directly attributable to performance-related factors. Moderate reserve levels, on the order of 2-3 months' imports for most countries, can help to buffer a good deal of exogenous volatility in aid. Since, on average, PRSC countries now hold close reserves close to 5 months import cover, some are already in a position where they can begin to use such an approach. Assuming that donors do indeed have longer aid horizons and that they are committed to following through, the success of this approach depends on two interrelated factors.

First, the reserve cushion works best when fluctuations in disbursements are serially independent or tend to offset each other in subsequent years. Such a pattern would be expected to result from administrative factors that cause uncertain delays. The cushion is most vulnerable when disbursement shocks are auto-correlated since countries will tend to experience a series of negative shocks over several years. Even in the worst cases, however, the size of the reserve provides a clear signal to donors over several years that disbursements are falling below

anticipated levels. This provides ample time for consultative groups to address the funding issue; the better they work, the lower can the reserve target be set.

Second, a clear performance framework will be needed to enable such a process to work smoothly. Donors are being asked, in effect, to fund a country's holding of reserves as insurance against shortcomings in disbursements that are not performance-driven. Lack of agreement on the performance framework will undermine such a system, as donors will see the fund as an escape valve for performance-related shortcomings; inadequate budget discipline in the country will also undermine the system. Mutually-agreed, independent, performance review, as pioneered in Tanzania and more recently extended to other countries, such as Mozambique, can strengthen consensus around the performance assessment.

In short, if donors are able to extend their horizon for aid funding, and they are really committed to following through their commitments with disbursements, and if a clear performance framework can be agreed, there is nothing standing in the way of solving the problem of short-term exogenous volatility in development assistance.

The second problem — whether it is possible to pre-commit aid for several years at a time without incurring high risks of misallocation — is a more complex one. While it can only be addressed quantitatively within a specific framework for aid effectiveness and allocation, most of the basic assumptions underlying the IDA-based model used here are common to many other approaches. Fully applying a performance-based allocation rule such as used for IDA can cut aid volatility considerably relative to historical levels, and pre-committing aid for several years ahead can further increase predictability. However, losses from pre-committing aid to all countries for several years can be substantial. Depending on assumptions on absorptive capacity, they could be over 10% of total aid for a five year period, with a plausible loss range between 5% and 21% of aid.

These losses can be cut dramatically using a flexible pre-commitment rule, where levels of assistance are revised sharply only in response to major changes in performance. Such a rule can increase predictability, especially for the more stable countries, while avoiding serious prolonged misallocations to “catastrophic” or rapidly improving countries. Indeed, a flexible commitment rule results in smaller losses, over time, than those which result from the likely measurement error in the performance ratings themselves. This makes it a very attractive option. Observed small performance changes frequently reverse themselves; possibly because they partly reflect measurement error. Another advantage of the flexible pre-commitment rule is that, by focusing on large and clearly observable, changes, it reinforces the credibility of performance-based allocation. Looking at the countries concerned, major performance changes within a short period of time very often seem linked to factors concerning political governance. Other studies show, and we assume, that aid in itself does not drive performance.

The third problem, of how to determine the share of budget support within an overall aid envelope, introduces a further degree of complexity. Budget support represents an “investment” of funds and reputation by donors to give clients the opportunity to develop stronger country systems of budget and financial management and service delivery. More so than for project aid, indicators of changes in country performance, rather than just levels of performance, are then

important indicators for justifying budget support. This can partly be taken into account by increasing the share of assistance provided through budget support as the level of performance increases. This approach would make budget support almost as predictable as the overall aid envelope. But if indicators are also considered necessary to signal the speed of change, the question is what they should be and how changes and levels should be weighted in determining disbursements.

Donors have different approaches to these questions. The World Bank, for example, has a tight, formula-driven method to determine the overall envelope but no formulaic approach to set either the desired level of budget support within the country program or the way in which levels of budget support will change in response to the fulfillment of prior actions. As a result, although in practice PRSC disbursements have been close to anticipated levels because the countries chosen for such operations have managed to sustain good track records, budget support from IDA is not explicitly predictable, except to the extent that countries higher in the ratings scale are more likely to receive it than those low down. The EU, on the other hand, uses a tight formula linking variable-tranche levels of budget support to indicators of effectiveness, including service delivery targets. Because the fixed tranche is usually large, the variability of EU budget support has been modest, about 8% of the mean level.

This paper does not take a position on whether prior actions or specific service indicators are better ways of conditioning changes in budget support levels on country “effort”. Indeed, it is interesting that the weights placed on public and financial management indicators (45%) and health and education indicators (25% and 22%) are not too different from the corresponding sectoral weights in PRSCs. We see the two approaches as complementary but imperfect ways to form a view on changes in the effectiveness of budget and financial management, service delivery systems, and in some cases, wider-ranging policies affecting private sector development and growth. But, unless these changes are “quantum” or “catastrophic”, the analysis of the CPIA shows that they cannot easily be observed from year to year. This argues for an approach closer to that of the EU, delineating a base level of support over a multi-year programmatic framework and using variable disbursement levels as incentive payments to encourage performance improvements. For overall allocations, IDA already uses such mechanisms including front-loading, back-loading and in more extreme cases, the high-low CAS scenarios. Similar approaches for PRSCs, holding incentive payments to around 10 percent of the base value, would enable disbursements to reflect trends in budget support that would be expected from positive or negative changes in CPIA scores over time.

This graduated approach would need to be complemented with three more processes. First, as for overall aid, budget support needs a reassessment trigger in response to major performance changes, equivalent to those that would result in a 0.3-0.4 point change in the CPIA, or an equivalent change in its governance and service-delivery-related components which are more critical for budget support.

Second, every three or so years requires an in-depth assessment of progress in the key areas, buttressed by output, financial management and service delivery indicators. This would both feed into the assessment that determine overall aid, and shape the decision on how much to provide in the form of budget support in the following multi-year period.

Third, historical information on the speed of progress can be used to help set output or outcome targets at challenging, yet reasonable, levels. This can both avoid setting countries up for the perception of failure by announcing unrealistic goals, and reduce the risk that countries will seek to maximize financial flows by setting goals that are easy to reach but reduce incentives.

As for overall aid, these approaches can help to make budget support more predictable and more credible, while containing the losses from large changes in performance.

One final issue is whether budget support can be used to strengthen the counter-cyclical potential of aid flows. The regular, annual cycle of disbursements makes it potentially useful, since funding can be increased or reduced in response to exogenous fiscal shocks. Natural disasters are often easy to identify at the country level, but this is less true for terms of trade and other worldwide changes which may be apparent at global level several months before they begin to be reflected in country-level data. To increase the feasibility of providing counter-cyclical support, we suggest adding a global monitoring variable to IDAs CPIA ratings for poor countries, which will become public after 2005. This would be the joint responsibility of the World Bank and the IMF, be updated quarterly, and indicate to donors whether global developments promised to deliver a negative, neutral, or positive shock to the country. This would cut down the information lag and mobilize support more quickly than in the past, with the objective of adjusting support levels to partially offset the impact of shocks on the budget. With such a system in place, the concessionality of compensatory financing would also be less of an issue, since donor grants could be used to pay off higher-cost finance more quickly.

## References

- Aizenmann, Joshua and Marion, Nancy (2004), "International reserve holdings with sovereign risk and costly tax collection," *The Economic Journal*, 114, 497: 569-91
- Aizenmann, Joshua; Lee, Yeonho; Rhee, Yeongseop (2004), ""International reserves management and capital mobility in a volatile world: policy considerations and a case study of Korea," UCSC Center for International Economics working paper 04-10
- Alesina, Alberto and Dollar, David (2000), "Who gives foreign aid to whom and why?" *Journal of Economic Growth*, 5: 33-64
- Andersen, Thomas; Hansen, Henrik; Markussen, Thomas (2004), "US politics and World Bank lending," available at <http://www.econ.ku.dk/okotand/images/IDA050804.pdf>
- Andersen, Thomas; Harr, Thomas; Tarp, Finn (2004), "US politics and IMF lending," discussion paper 04-11, Institute of Economics, University of Copenhagen
- Baldacci, Emanuele, Clements, Benedict, Gupta, Sanjeev, and Cui, Qiang (2004) "Social Spending, Human Capital and Growth in Developing Countries: Implications for Achieving the MDGs" IMF Working Paper 217. November
- Bulir, Ales and Hamann, Javier (2003), "Aid volatility: an empirical assessment," *IMF Staff Papers*, 50, 1: 64-88
- Bulir, Ales and Hamann, Javier (2005), "Volatility of Development Aid: From the Frying Pan into the Fire?" working paper, IMF
- Burnside, Craig and Dollar, David (2000), "Aid, policies and growth," *The American Economic Review*, 90, 4
- Cashin, Liang and McDermott (2000), "How persistent are shocks to world commodity prices?" *IMF Staff Papers*, 47, 2: 177-217
- Clemens Michael A. (2004), "The Long Walk to School: International Education Goals in Historical Perspective," Center for Global Development Working Paper # 37
- Clemens Michael A, Radelet Steven and Bhavani Rikhil (2004), "Counting Chickens when they Hatch: The short-term effect of Aid on Growth", Center for Global development Working Paper #44
- Clemens, Michael A; Kenny, Charles A; Moss, Todd J (2004), "The Trouble with the MDGs: Confronting Expectations of Aid and Development Success," Center for Global Development Working Paper #40

Collier, Paul and Dehn, Jan (2001), "Aid, shocks and growth," World Bank policy research working paper #2688

Collier, Paul and Dollar, David (2002), "Aid allocation and poverty reduction," *European Economic Review* 45, 1

Collier, Paul (1999), "Aid dependency: a critique," *Journal of African Economies*, 8, 4: 528-45

Denkabe, Philip (2003), "Policy, aid and growth: a threshold hypothesis," Center for Global Development seminar paper, <http://www.cgdev.org/Events/Index.cfm?EventID=66>

Department for International Development, Great Britain (2005), "Financing Poverty Reduction Strategies to Meet the MGDs – Some Process Issues," draft

Dollar, David and Levin, Victoria (2004), "The increasing selectivity of foreign aid, 1984-2002," World Bank Policy Research Working Paper #3299

Edwards, Sebastian and van Wijnbergen, Sweder (1989), "Disequilibrium and structural adjustment," in Chenery and Srinivasan (eds.), *Handbook of Development Economics*, Vol. II. Amsterdam: North-Holland.

European Commission (2005), "EC Budget Support: An Innovative Approach to Conditionality and Development," February draft.

Foster, Mick (2005), "Improving the medium-long-run predictability of aid," draft paper prepared for DFID.

Gabriele, A.; Baratav, K.; Parikh, A. (2000), "Instability and volatility of flows to developing countries," *The World Economy*, 23, 8: 1030-56

Gelb, Alan; Ngo, Brian; Ye, Xiao (2004), "Implementing Performance-Based Aid in Africa," World Bank, Africa Region Working Paper #77

Hansen, Henrik and Tarp, Finn (2000), "Aid effectiveness disputed," *Journal of International Development*, 12 (April): 375-388

Hansen, Henrik and Tarp, Finn (2001), "Aid and growth regressions," *Journal of Development Economics*, 64 (April): 547-70

Hermann, R.; Burger, K.; Smit, H.P. (1990), "Commodity policy: price stabilization versus financing," in Sapford and Winters (eds.), *Primary Commodity Prices: Economic Models and Policy*, Cambridge University Press

Koenker, Roger and Basset, Gilbert (1978), "Regression quantiles," *Econometrica*, 46, 1: 33-50

Kostopoulos, Christos (1999), "Progress in Public Expenditure Management in Africa: Evidence from World Bank Surveys," ARWPS - 1

Lensink, Robert and Morrissey, Oliver (2000), "Aid instability as a measure of uncertainty and the positive impact of aid on growth," *Journal of Development Studies*, 36, 3

Neumayer, Eric (2003a), *The Pattern of Aid Giving*, London: Routledge.

Neumayer, Eric (2003b), "The determinants of aid allocation by regional multilateral development banks and the United Nations agencies," *International Studies Quarterly*, 47: 101-122

OECD/DAC (2005), Paris Declaration on Aid Effectiveness, DCD/DAC/EFF, February.

OECD/DAC (2004), Draft Good Practice in Public Finance Management: Increasing the Predictability of Aid Flows. Draft, September.

Osei, Robert; Morrissey, Olivier; Lensink, Robert (2002), "The volatility of capital inflows: measures and trends for developing countries," CREDIT research paper no. 02/20

Pallage, Stephane and Robe, Michael (2003) "On the Welfare Cost of Economic Fluctuations in Developing Countries," *International Economic Review*, vol. 44(2), pages 677-698

Pallage, Stephane; Robe, Michael; Bérubé, Catherine (2004), "On the Potential of Foreign Aid as Insurance," Research paper 04/04, CIRPEE.

Pallage, Stephane and Robe, Michael (2003). "Leland & Pyle Meet Foreign Aid? Adverse Selection and the Procyclicality of Financial Aid Flows," Research paper 03/27, CIRPEE.

Perry, Guillermo (2003), "Can Fiscal Rules Help Reduce Macroeconomic Volatility in the Latin America and Caribbean Region?" World Bank Policy Research Working Paper # 3080

Strategic Partnership for Africa (2005), Budget Support Working Group Report to Plenary, January

Varangis, Panos; Varma, Sona; dePlaa, Angelique; Nehru, Vikram (2004), "Exogenous shocks in low-income countries: economic policy issues and the role of the international community," mimeo, World Bank.

Varangis, Panayotis and Larson, Donald (1996), "Dealing with commodity price uncertainty," World Bank policy research working paper #1

World Bank 2004, Strategic Framework for IDA in Africa

World Bank 2005 "A Stocktaking of PRSCs". OPCS, Draft March 7